The Johns Hopkins University
APPLIED PHYSICS LABORATORY
Silver Spring, Maryland

(14) Rept. no. TG-653

~~Copy 032~~

(5) Applied Physics Lab., Johns Hopkins Univ., Silver Spring, Md.

Jan[redacted]65, (11)

(6) ANALYSIS OF THE OBSERVATIONAL CONTRIBUTIONS TO THE ERRORS OF THE NAVY SATELLITE DOPPLER GEODETIC SYSTEM*,

(10) by W. H. Guier, R. R. Newton, and G. C. Weiffenbach.

An analysis is made ~~of~~

Abstract

~~This report analyzes~~ the observational contributions to the total system error of the Navy's Satellite Doppler Geodetic System. Aside from the error imposed by the limited number of satellites available for analysis, which limits the complexity of the gravity field that can be used, it is concluded that the system error is about 10 meters for the observations from a single satellite pass, and is 5 meters or less for multiple passes.

The figures, tables, and references applying to each sub-section will be found at the end of that sub-section.

*This work was supported by the United States Department of the Navy, Bureau of Naval Weapons, under Contract NOw 62-0604-c. (15)

21739

mB

Foreword

The persons whose names appear on the title page of this report have placed them there to indicate that they accept responsibility for its contents, and are not the sole contributors. Helen S. Hopfield contributed most of sub-section III.1, M. M. Feen most of sub-section III.2, and S. M. Yionoulis the material on resonance in sub-section IV.2. All contributors are staff members of the Applied Physics Laboratory.

ANALYSIS OF THE OBSERVATIONAL CONTRIBUTIONS TO THE ERRORS OF THE NAVY SATELLITE DOPPLER GEODETIC SYSTEM

I. INTRODUCTION AND SUMMARY

Beginning in 1959 and continuing to the present, the Navy has built up and maintained a system and a program for geodetic research using measurements of the Doppler shift in radio receptions from artificial satellites. The ground system involved is frequently called the TRANET system, and the program is frequently the ANNA program. Technical responsibility for the ground system, including communications, and for the satellites, has been placed at the Applied Physics Laboratory of The Johns Hopkins University. Geodetic analysis using only data from this system has also been carried out at the Applied Physics Laboratory. Analysis using data from this system as well as other data has been performed at the Naval Weapons Laboratory, Dahlgren, Virginia.

Data from the system can be used either for geometrical geodesy or for dynamical geodesy. In geometrical geodesy, the objective is to find the magnitude and direction of the vector joining two identifiable points on the Earth. In dynamical geodesy, the objective is to determine the magnitude and direction of the acceleration due to gravity at any point on the Earth; a small amount of geometrical information is necessarily obtained in the process.

So far, data from the system have been used almost entirely for dynamical geodesy, although there is a growing interest in use for geometrical geodesy.

Errors in the final results obtained after the analysis depend upon many factors. These factors can be loosely put into two groups. We shall call one of these groups "observational", defined roughly as containing all of the factors that affect the quantity and quality of the raw data which are the inputs to the analysis program. We shall call the other group "analytical", defined as including those phenomena that affect the analytic results but which are not included in the analysis for some reason.

We believe that there is one analytical error that completely dominates all other errors, whether observational or analytical, in the present dynamical results of the system. The error arises from neglected components in the gravity field: With a given amount of data, it is possible to estimate only a limited number of gravity components; this limitation will be discussed in Section IV. The neglected components produce orbital effects having some similarity to effects of the components estimated in the analysis. The analysis program finds, for those components that it tries to estimate, the values that minimize the RMS error. Any estimated component thus includes two parts. One part is the "true" value. The other part is the value of the component which does the best job of masquerading as one or more neglected components. The mask can be stripped from this latter part only by obtaining more data.

Because of this state of affairs, we believe that it is best to deal with observational and analytical errors in separate reports. This report will deal with the observational factors. Thus it will remain valid even if large advances in the analytic program decrease the total system error dramatically. Of course, improvements in the observational errors can still occur, but for a long time they will be substantially independent of the analytical errors or of the total error. The analytical errors are discussed in Ref. 1.

In order to discuss errors quantitatively, it is necessary to have a meaningful measure of error. For geometrical geodesy, there is an obvious measure, namely the error in a measured position, or rather, the estimated world-wide RMS error in a measured position. It is fairly easy to discuss the observational errors in terms of the position error that they introduce.

For dynamical geodesy, one could use the estimate of the RMS error in the gravity vector for all points on the surface of the Earth. Unfortunately it is not easy to relate many types of system error to this measure. Thus even for dynamical geodesy, we shall use a measure of error that is directly derived from geometrical geodesy. Assume that the orbit of a satellite is known exactly, and that the data from a single pass of the satellite are used to derive the position of the observer. The measure of error used in this report will be the best estimate of the standard deviation of this derived position. This measure is called "tracking residual" in Ref. 1.

In this report, we conclude that the contribution made by observational factors to total system error is around 10 meters. By contrast, the total system error found in Ref. 1 is around 75 meters, and therefore observational factors do not contribute directly any appreciable part of the total error.

There is an indirect contribution of what we are calling observational factors in this report. Among these factors is the number of different orbits from which data are available; it is arbitrary whether we call this factor observational or analytical. This number limits the number of gravity components that can be estimated, and therefore determines the components which must be neglected. Since we believe that the neglected components are the principal source of error, there is a relation between system error and the observational factor of the available orbits.

If the system is used for geometrical geodesy, the errors depend somewhat upon the length of the vector being determined. Over a short distance (say 100 km. or less) the error is probably only that arising from the instrumentation. Over larger distances, say up to 1000 km. or so, additional error probably arises from refraction effects which may now be somewhat different at two locations whereas they were almost identical at nearby sites. Over still longer distances, analytical errors begin to contribute appreciably: The greater the distance, the greater the effect of an error in the assumed satellite orbit upon inferred relative position.

For short distances, the geometrical error is about 10 meters for a single pass and 1 or 2 meters for a reasonable number of passes analyzed together. Over the dimensions of the continental United States, these figures rise to about 50 meters and 10 meters respectively.

In Section II we shall discuss the errors arising from the instrumentation. In Section III we shall discuss errors arising from propagation problems between the satellites and the ground. Finally, in Section IV we shall discuss the limitations imposed upon the results obtained from the system by the configuration of ground stations and by the configuration of the satellite orbits used.

The error analysis of this report and of Ref. 1 discloses the following sources of error in the geodetic system:

	Meters
Random measurement error:	8
Timing error:	3
Uncorrected tropospheric refraction:	3
Uncorrected ionospheric refraction:	3
Effects of neglected gravity harmonics:	75
Resultant system error for dynamical geodesy:	76

As we have already discussed, the system error for geometrical geodesy is much less than this, and is probably about 10 meters over an area the size of the continental United States.

As the effects of neglected gravity harmonics are reduced by determining more harmonic coefficients, requiring more satellites, this contribution will approach zero. The system limit is presumably that imposed by the other error sources. With no improvement in techniques, this limit is $(8^2 + 3^2 + 3^2 + 3^2)^{\frac{1}{2}} = 9.5$ meters for the observations from a single pass. With enough passes to reduce random errors to a negligible amount, we are left with the bias errors. The timing and the refraction errors tend to give a bias for a particular geometry of a pass. Averaged over all passes, they have some tendency to cancel but probably have some irreducible bias of as yet unknown amount, particularly the uncorrected ionospheric refraction. Thus the system limit with present techniques is somewhere between 0 meters and $(3^2 + 3^2 + 3^2)^{\frac{1}{2}} = 5.2$ meters. For the present, we shall adopt 5 meters as the best estimate of the observational error for multiple passes.

The observational errors have been larger in the past. A study of the random measurement error in Ref. 2 using data from early 1962 gave 28 meters instead of 8, and it is possible that some of the tracking equipment still in use has a measurement error that large. Timing errors before late 1962 when the first satellite timing became available may have contributed the order of 15 meters for some stations. Data collected since satellite timing became available may still have this much error in the raw data, but are correctable to a level of 3 meters or less. Finally, the uncorrected ionospheric refraction error may be much larger, at maximum solar activity, perhaps several tens of meters, unless suitable measures can be taken.

References - Section I.

1. Guier, W. H., and Newton, R. R., Status of the Geodetic Analysis Program at the Applied Physics Laboratory, Confidential, Applied Physics Lab. Report TG-652, January 1965.

2. Newton, R. R., Errors in Long-Term Orbital Prediction for Satellite 1961 o 1, J. Geophys. Res., 69, pp. 3619-24, September 1, 1964.

II. ACCURACY OF TRANET TIME AND FREQUENCY MEASUREMENTS

Instrumentation errors in TRANET doppler measurements can be classified in two categories, random and systematic, where random is meant to indicate those errors which produce position errors that tend to cancel when passes from some given satellite are combined to form an average position. Systematic errors are here defined to be those error contributions which will produce calculated station positions that are systematically biased in some particular direction for all passes from some one satellite. Instrumentation errors can be further classified as timing (epoch) and frequency errors.

For any single satellite pass the position errors attributable to TRANET frequency measurements are primarily random errors of 10 meters or less. This is shown in Fig. II-1, which is reproduced from Ref. 1. To obtain this figure, two tracking stations located at the same site obtained data simultaneously on each of about fifteen satellite passes. A position was then inferred for each station. Each point in the figure is the difference between the inferred positions for a single pass. Under these circumstances, systematic errors, including errors in the satellite orbit used for deriving station positions, should cancel amost exactly, leaving only the random errors in measurement as the sources of the error.

The standard deviation of the difference in position is 5.0 meters in the north-south direction and 10.4 meters in the east-west direction. Since each point is the result of two independent measurements, the measurement error for a single pass should be $5.0/\sqrt{2} = 3.8$ meters in one direction and $10.4/\sqrt{2} = 7.4$ meters in the other, for a resultant of 8.1 meters. For general discussion, we shall use 10 meters for the random component.

While errors due to frequency measurement tend to be random, errors attributable to TRANET timing inaccuracies tend to be systematic, and at the time of this writing can be as large as 15 meters at some stations. The use of satellite timing, which is now being implemented, will reduce this source of error to 3 meters or less.

A brief description of the doppler measuring procedures used in the TRANET system will serve to elucidate the various sources of error. The block diagram in Fig. II-2 illustrates the system components pertinent to this discussion.

The satellite contains an ultra-stable quartz crystal oscillator mounted in a multiple Dewar flask. The quartz crystal, in most cases a 5 mc fifth overtone AT cut unit mounted in an evacuated glass envelope, is embedded in an 8 oz. monel cylinder which provides a large heat capacity. The monel cylinder is then wrapped in alternate layers of aluminum foil and fiberglass paper, which provide both low thermal conductivity and low radiative heat transfer, and this assembly in turn is mounted in an intermediate aluminum container. A heater coil and a mercury-in-glass thermostat are mounted in good thermal contact to this aluminum cylinder and cycle the latter over a range of $0^{\circ}.1$C in the vicinity of the crystal turnover (zero temperature coefficient) point. The large thermal time constant (about 10 hours) between the intermediate cylinder and the crystal reduces the short-term temperature fluctuations of the crystal to less than 0.001°C. To reduce the power needed in the heater coil and to add more thermal isolation the crystal oven is enclosed in additional layers of insulation and mounted in an outer aluminum can. Two buffer amplifier stages and a Zener diode voltage regulator are also mounted inside the oven for added stability. The power and signal leads into the oscillator are No. 32 nichrome wire to provide high thermal resistance. The entire oven assembly is then mounted in the satellite in a location selected for minimum temperature fluctuations, and provided with regulated power.

These oscillators have demonstrated good stability both in the laboratory where extensive tests are run on all units, and in orbit. A typical case is an oscillator for the GEOS A satellite which is currently under test and for which the following data were obtained:

averaging time	RMS stability
2 sec	12×10^{-12}
20 sec	5×10^{-12}
200 sec	7×10^{-12}
2000 sec	13×10^{-12}

The long-term drift of these oscillators is essentially the aging rate of the quartz crystals, and typically lies between 2×10^{-10} and 2×10^{-11} per day.

The stable oscillator output drives the doppler transmitters through multiplier chains, and the satellite clock (if used*) through

---

*Satellite clocks are not required in all doppler geodetic satellites since the doppler system requires timing only to synchronize station clocks.

a divider chain. High phase stability is maintained in transmitters and multipliers by the use of high Q circuits and tight phase comparison and feedback circuits. The phase stability required in the divider circuits is relatively modest (some tens of microseconds) and can be achieved without difficulty.

Modulation if used for timing or telemetering can be of two types, continuous or intermittent, but is always a square wave phase modulation generated by 3-state switches. When the modulation is continuous, each bit consists of a balanced pattern of phase advance, an accurately matched phase retard, and normal, with the modulation frequency selected to generate only those sideband frequencies which can be readily separated from the doppler carrier in the ground station. Intermittent modulation, such as used in ANNA IB, is comprised of alternate cycles of phase advance and retard with a return to normal phase after each burst, the modulation typically being kept on for 1/3 second and off (normal phase) for a minute or longer. In both types of modulation the doppler carriers retain the requisite high phase stability, and the entire modulation pattern is controlled by the stable oscillator in a precisely known fashion. The actual fiducial time marker is either a specified pattern of bits ("barker" word) or a phase reversal in the modulation pattern. In either case the fiducial time marker is applied periodically once per minute, or some comparable interval.

The block diagram for the TRANET ground stations shown in Fig. II-2 has been simplified, for example, by omitting the dual-frequency refraction correction, but illustrates those elements of concern to the discussion of this section. In addition, the tracking filter shown as a separate box in the diagram is incorporated as part of the receiver in several TRANET stations, the receiver then being designated a "phase-lock receiver". Nonetheless, all of the present TRANET stations are properly described in a functional sense by Fig. II-2.

From the standpoint of error contributions the receiver as such need only be considered in terms of its noise figure and time delays. Its function is to select the appropriate satellite signals, subtract these signals from a reference frequency generated in the station, and amplify them for presentation to the phase-locked tracking loop, or tracking filter. The satellite transmitter power outputs that are available are such that the receiver noise figures that can be attained (10 db or less) introduce no significant errors. Furthermore, errors from this source are truly random in the sense described above. Time, or phase, delays in the receivers can readily be kept to insignificant levels since all of the pertinent circuits are wide-band (100 kc or greater) tuned filters*. Total receiver delays are typically less than 10 microseconds.

The phase-locked tracking loops need more detailed discussion, as they do contain very narrow bandwidth low-pass filters. A block diagram is shown in Fig. II-3. When phase-lock is maintained, the VCO (voltage controlled oscillator) frequency is displaced from the satellite signal out of the receiver by an amount equal to f_1, the tracking filter's stable

---

*It should be noted that there is no requirement in doppler tracking to maintain very precise absolute phase throughout the system as is the case in range measurements. Here it is only necessary to conserve the epoch of the doppler measurements to about 0.5 milliseconds.

oscillator frequency. The mixer output is the difference frequency, or f_1, which is then amplified in the tuned IF amplifier (whose center frequency is also f_1). The phase detector receives the amplified difference frequency, phase compares it with f_1 as generated by the stable oscillator, and generates a DC error signal which is a function of the relative phases of the two phase detector inputs. The error signal then goes through the narrow band low-pass filter which removes virtually all noise, is amplified and applied to the VCO. The sense of the error signal is such that if the VCO-input signal difference deviates from f_1, that is if the two inputs are not identical in phase, the VCO frequency will be driven by the error signal to maintain the difference f_1 at the input mixer at an in-phase condition at the phase detector. The VCO output should then be a smoothed replica of the input signal, but displaced by the frequency f_1. A second mixer is then employed to subtract f_1 from the VCO frequency, providing an output which is identical to the desired signal in frequency but with a greatly reduced noise level. The effective bandwidths of the tracking filters range from 1 cps to 50 cps.

The primary reason for employing such a complex filter is that one can apply very narrow bandwidths to a maneuvering doppler signal without incurring appreciable phase delays. This can be seen by noting that if perfect phase-lock can be maintained, i.e., if the two inputs to the phase detector are kept exactly in phase by the loop, the phase difference between the slowly varying doppler input and output signals is only that which is introduced by the output mixer circuitry. Since no significant additional filtering is needed in the output mixer this difference can be very small. In practice, of course, this ideal condition is never realized since perfect phase-lock is never quite achieved. Nonetheless, if adequate signal is received to maintain phase-lock at all, none of the narrow-band elements contribute a phase delay. The maximum phase shift that occurs under phase-lock conditions is 1 radian - a prerequisite to maintaining lock - at the frequency f_1, plus the phase delay in the IF amplifier. The latter can be made quite small, since the IF frequency selective elements are tuned filters which have zero phase shift at their center frequencies.

The various tracking loops employed in TRANET stations differ in some details from the circuit described above, but the only material differences for our present purpose are the use of additional filters at various points to suppress spurious frequencies, and the detuning action of AGC circuits on the tracking-loop IF amplifiers. These can be monitored and calibrated. Typically the nominal phase shift from receiver input to the tracking-loop output lies in the range ± 500 microseconds. In any given set of equipment long-term variations from the nominal value can be readily maintained at less than 100 microseconds, and with somewhat more meticulous alignment and calibration can be controlled to better than 25 microseconds. Variations within the time of a pass (primarily due to AGC action and thus dependent on signal amplitude) can be as large as 200 microseconds in some equipments.

The frequency standards in the stations are used to generate the reference frequencies which are mixed with the satellite signals in the receivers, and thus contribute directly to the doppler frequency errors.

These same standards are also used to control the station clocks and to provide the meter frequency and timing inputs to the doppler frequency digitizer. The standards used are high quality commercial units with frequency stabilities (measured in the APL Time and Frequency Laboratory) as follows:

averaging time	RMS frequency stability
1 sec	20 to 50×10^{-12}
10 sec	5 to 10×10^{-12}
100 sec	$1\ 1/2$ to 4×10^{-12}
1000 sec	1 to 3×10^{-12}
1 day	approx. linear drift of 2×10^{-11} to 2×10^{-10}

All TRANET stations maintain a continuous monitor of these standards by means of standard frequency VLF transmissions, taking into account the diurnal phase shift which occurs in all VLF transmission. Each VLF station used for this purpose is also monitored in the APL Time and Frequency Laboratory.

The latter facility maintains two Cesium Beam frequency standards and several quartz crystal standards, all of which are monitored with respect to each other and with respect to WWV via VHF transmissions and the National Frequency Standard in Boulder, Colorado via VLF. An accurate epoch is maintained in this laboratory with respect to both WWV, again using VHF, and relative to Universal Time as maintained by the U. S. Naval Observatory. This time synchronization is facilitated by the proximity of the WWV transmitter (about 10 miles) which minimizes uncertainties in propagation time. Synchronization has also been checked by carrying a precision quartz crystal portable clock from APL to WWV and to the Naval Observatory. It is possible by these means to maintain the APL time and frequency standards to within 10 microseconds or better in epoch and to about 2×10^{-11} in frequency relative to the primary standards. (TRANET station 111 is located in the same building as the APL Time and Frequency Laboratory and has direct access to these standards. Thus it is the station used to calibrate our satellite clocks.)

The last section of the station block diagram that we need to consider is the analog to digital converter. It is the function of this unit to convert the received doppler frequency as presented to it by the tracking filter to digital form and reference these data to the station clock. Here again the various stations differ in detail, but are functionally identical. The block diagram of Fig. II-4 illustrates the method used. A clock pulse from the station clock opens gate 1 and permits the doppler signal to enter the preset counter. The events detected by this counter are the positive-going zero crossings of the doppler signal. It puts out a start signal to gate 2 on the first of these events, counts each successive event until it has registered a total of n_c events, or cycles, and then sends a stop signal to gate 2. Gate 2 is thus open for an interval corresponding to n_c doppler cycles. A "meter frequency" from the frequency standard is sent through this gate to the period counter. Since the meter frequency has an accurately known period, the period counter provides an accurate measure of the average period of the n_c doppler cycles.

In practice, n_c is a fixed number for an entire satellite pass, and is selected as the largest number that can be used without having the period counter read as high as 1 second for the lowest expected doppler frequency. The meter frequency is either 1 mc or 5 mc, providing resolutions of 1 microsecond and 0.2 microsecond respectively. It is possible to make a measurement once each 2, 4, 8 or 16 seconds throughout a pass, but 4 seconds is the usual interval. Each measurement is combined with the time of the UT second marker pulse used to open gate 1, and punched on paper tape for transmission to the computer.

The only significant error introduced by this digitizing process is the ± 1 count (± 1 or ± 0.2 microseconds) ambiguity in the period count, which results in a frequency error of $\pm (f^2/n_c) \times \tau$ cps, where f is the doppler input to the digitizer and τ is the meter frequency period, typically about 10^{-10} (RMS) of the doppler transmitter frequency. This error is clearly random.

At the present time all but one of the TRANET stations employ standard time broadcasts to synchronize their clocks. (The one exception is station 012 in Australia which uses satellite timing.) With the exception of station 111 at APL, this procedure introduces significant systematic errors, ranging from 0.5 to 2.5 milliseconds. The use of satellite time synchronization will result in overall timing errors of 0.2 to 0.4 millisecond. All other sources of timing error are small, such as uncertainties in time delays in the receiving system for both doppler and timing signals, or small and random, for example, jitter in satellite and station clocks (~ 1 microsecond or less) and fluctuations in the interval between the opening of gates 1 and 2 in the digitizer (50 microseconds or less). In sum, the timing errors attributable to instrumentation are primarily station clock synchronization errors, and these will result in (possibly) systematic errors of the order of δt x satellite speed, or 4 to 18 meters along-track error at present, reducible to 1.5 to 3 meters with satellite timing.

The frequency errors attributable to instrumentation produce no significant systematic effects, since the major sources of frequency errors, namely the satellite oscillators, station frequency standards, and the doppler digitizer, are either completely random as in the case of the last mentioned or have variations which are almost completely uncorrelated from pass to pass*. This can be seen in the following table which lists the various expected errors using a conversion from frequency error to equivalent position error which represents an absolute upper limit for the worst geometry and other conditions.

Source of Error	Averaging Time	RMS Freq. Error	Maximum Resulting Position Error
Satellite Oscillator	2 sec	12×10^{-12}	0.1 meter
	20 sec	5×10^{-12}	0.1 meter
	200 sec	7×10^{-12}	0.4 meter
	2000 sec	13×10^{-12}	6 meters
	Long Term Drift	$\sim 2 \times 10^{-10}$/day	0.1 meter

*Smooth drift in the station and satellite oscillators is calculated and corrected in the analysis programs.

Source of Error	Averaging Time	RMS Freq. Error	Maximum Resulting Position Error
Station Freq. Std.	1 sec	5×10^{-11}	0.5 meter
	10 sec	10^{-11}	0.2 meter
	100 sec	4×10^{-12}	0.2 meter
	1000 sec	3×10^{-12}	1.5 meter
	Long Term Drift	2×10^{-10}/day	0.1 meter
Digitizer	$\sim$1 sec	10^{-10}	1 meter

Since changes in equipment can influence the above estimates, and because there are pertinent differences among the various TRANET stations, Table II-1 lists the equipment currently used in each location. In addition, the antenna type and separation is also listed for each station, since these can also introduce geodetically important errors.

References - Section II.

1. Newton, R. R., Measurements of the Doppler Shift in Satellite Transmissions and Their Use in Geometrical Geodesy, Applied Physics Lab. Report TG-627, November 1964.

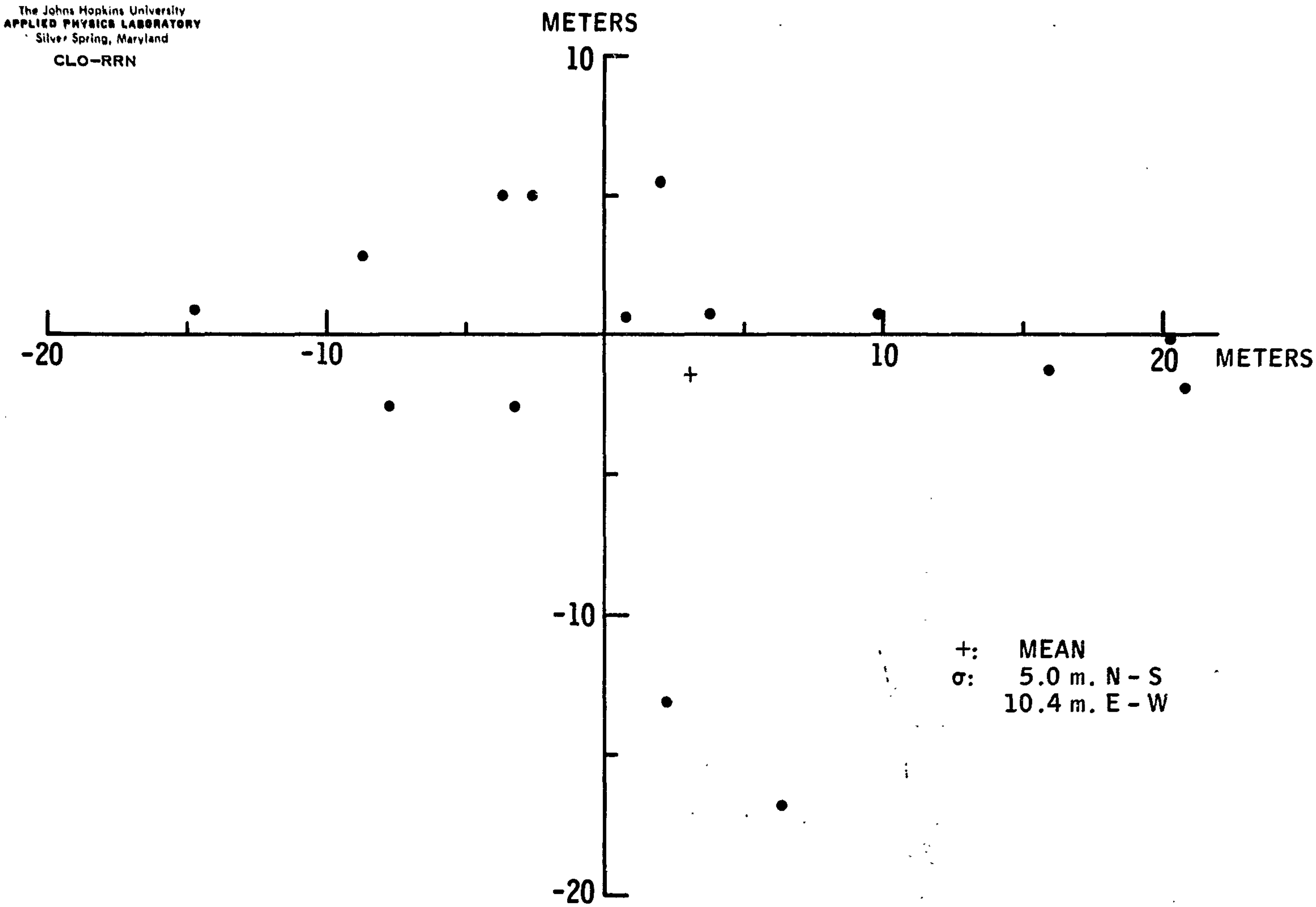

Fig. II–1. DIFFERENCES OF STATION COORDINATES OBTAINED FROM TWO COLLOCATED DOPPLER TRACKING STATIONS.

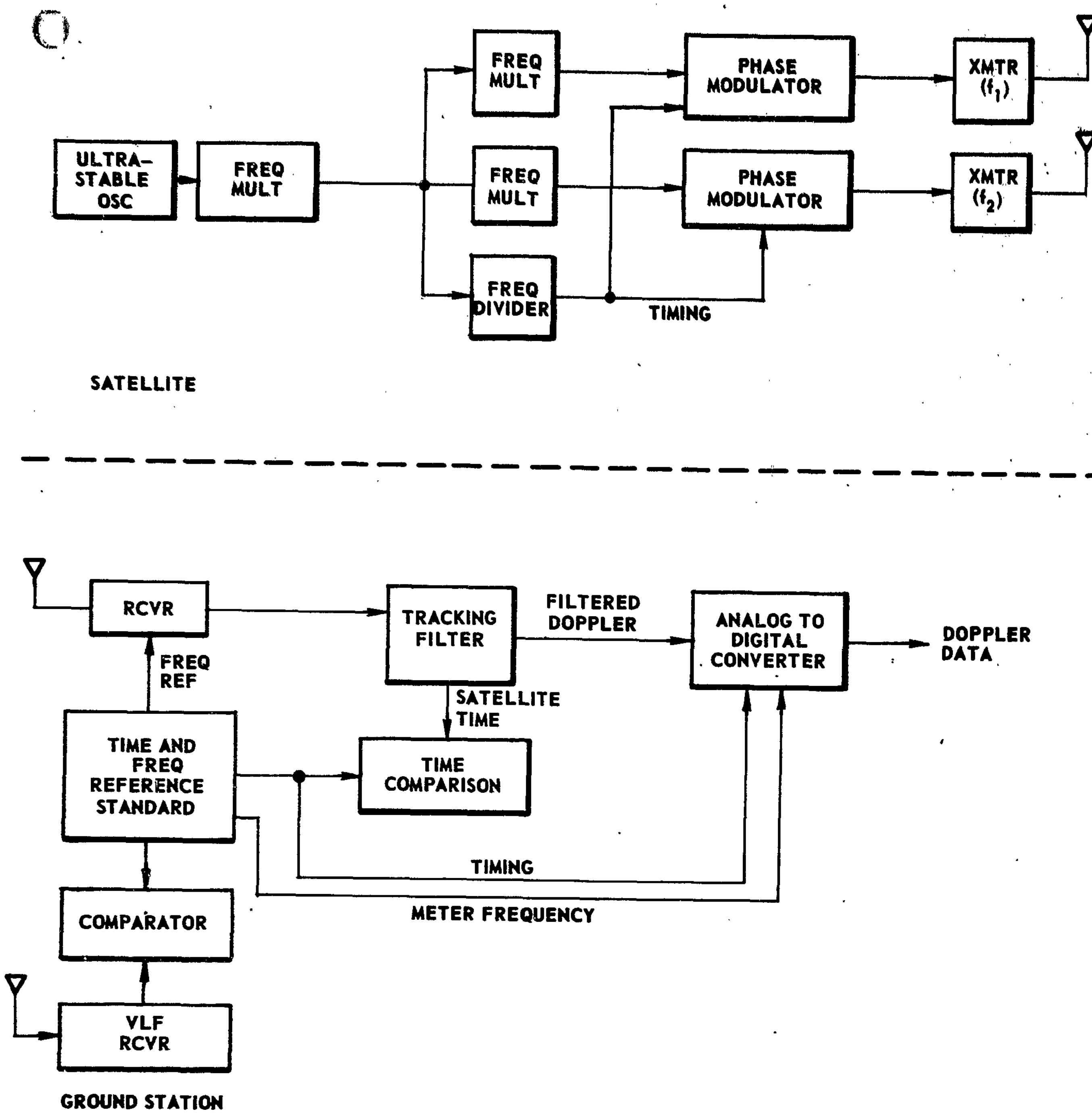

Fig. II-2 BLOCK DIAGRAMS OF FREQUENCY AND TIMING EQUIPMENT IN DOPPLER GEODETIC SATELLITES AND TRANET GROUND STATIONS.

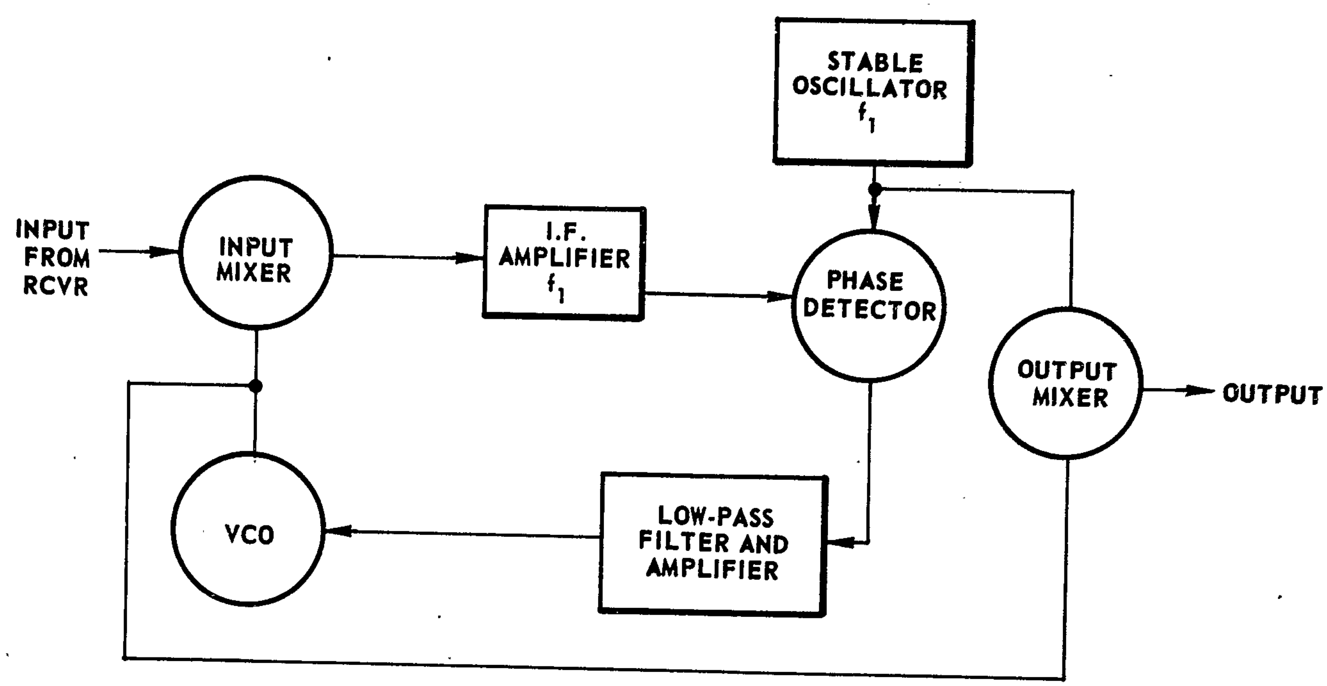

Fig. II-3 BLOCK DIAGRAM OF A PHASE-LOCKED TRACKING LOOP.

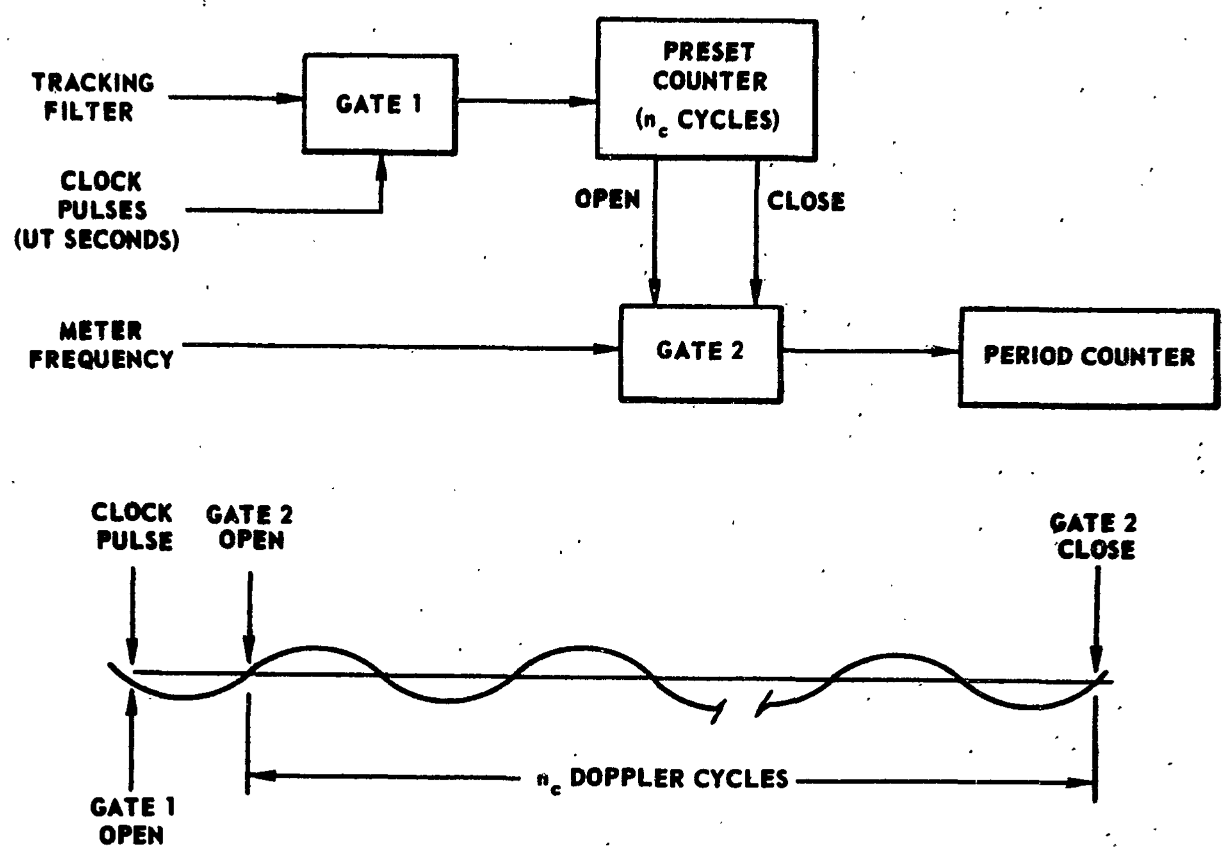

Fig. II-4 DOPPLER FREQUENCY DIGITIZER.

Table II-1

Characteristics of TRANET Tracking Stations

STATION LOCATION	TRACKING FILTER	METER FREQUENCY	FREQUENCY STANDARD	STATION CLOCK REFERENCES		Current Probable Errors (MS)	Antenna Type	SEPARATION (meters)	
Fixed Stations				Epoch	Clock Rate			Within Pairs	Between Pairs
003 - Las Cruces	1-Interstate M VIII	1-MC	Sulzer 2.5	WWV	* NPG	1.5	Whip	4	3
008 - Brazil	Space General	5-MC	Hermes 105	WWV	* NPG	1	Whip	2	1
011 - Philippines	Space General	5-MC	Hermes 105	WWVH	* NPG	1.9	Whip	?	?
012 - Australia	ITT	5-MC	Hermes 105	Satellite	* NPG	.2	Whip	13	5
013 - Japan	ITT	5-MC	Hermes 105	JJY	* NPG	2.4	Whip	2	2
014 - Alaska	2-Interstate M VIII	1-MC	Sulzer 2.5	WWVH	* NPG	1	Whip	3	2 1/2
017 - Samoa	Space General	5-MC	Hermes 105	WWVH	* NPG	.4	Whip	2	2
018 - Thule	ITT	5-MC	Hermes 105	WWV	* NPG	.5	Whip	2	1
019 - McMurdo	ITT	5-MC	Hermes 105				Whip	-	-
092 - Austin, Texas	2-Interstate M VIII	1-MC	Hermes 101C	WWV	* WWVL	.4	150-160 7 element Crossed Yagi 324-400 12 element Crossed Yagi	?	?
106 - England	2-Interstate M VIII	1-MC	Varian Cesium	MSF	GBR	1.4	Whips - 54-150-400 -162 Yagis - 162-324	?	?
111 - How. Co.	3-Interstate M VIII	1-MC	Sulzer 2.5	WWV	APL Stds Lab	.2	Whip	2	2
115 - South Africa	ITT	5-MC	Hermes 105	ZUO	* GBR	.5	Whip	?	?
717 - Seychelles	2-Electrac 215	5-MC	Sulzer 2.5	WWV	* NPM	?	Whip		

Table II-1 cont'd.

STATION LOCATION	TRACKING FILTER	METER FREQUENCY	FREQUENCY STANDARD	STATION CLOCK REFERENCES		Current Probable Errors (MS)	Antenna Type	SEPARATION (meters Within Pairs	Between Pairs
Van Stations				Epoch	Clock Rate				
306 (V-4)	2-Electrac 215	5-MC	Sulzer 2.5,5	*	VLF				
307 (V-3)	2-Electrac 215	5-MC	Sulzer 2.5,5	*	VLF				
311 (V-1)	2-Interstate M VIII	5-MC	Sulzer 2.5,5	*	VLF				
312 (V-2)	2-Interstate M VIII	5-MC	Sulzer 2.5,5	*	VLF				
314 (V-5)	2-Electrac 215	5-MC	Sulzer 2.5,5	*	VLF				

* Stations are now or will shortly be equipped with the PSL VLF Clock and Standby Batteries which, properly used, should make it possible to maintain epoch to within ± 30 μsec. with respect to the transmitted VLF signal.

III. REFRACTION CONTRIBUTIONS TO THE ERROR

When satellite signals are propagated through the earth's atmosphere, errors are introduced in the data as received at the antenna of the receiving station and are essentially uncorrelated with station instrumentation errors. In the frequency regime of present interest there are two classes of errors that are known to be non-negligible at least part of the time. These errors are usually termed refraction errors and fall into two classes whose characteristics and basic sources are different. The first is tropospheric refraction error which is caused by the lower part of the atmosphere associated with the earth's weather. The second is ionospheric refraction error caused by the propagation of the signals through the uppermost layers of the atmosphere. These upper layers are ionized by the sun's radiation and form a plasma known as the ionosphere. The characteristics of these two errors and the methods for correcting their effects are considered separately.

III.1. Tropospheric Refraction Errors

An excellent analogy exists between the effects produced by the troposphere on UHF transmissions and the optical refraction that occurs in optically refractive media. For this reason, the refraction effects caused by the troposphere can be considered in terms of a refractive index, $\underline{n}$, with the propagation obeying Fermat's Principle. Consequently, the portion of the phase path which is in the troposphere, $s = \int n \, ds$, differs from the corresponding portion of the instantaneous slant range. The difference, called Δs_{tro}, is not constant during a satellite pass, being obviously greater at low than at high elevation angles. The tropospheric contribution to the Doppler shift at any instant is then

$$\Delta f_{tro} = - \frac{f}{c} \frac{d}{dt} (\Delta s_{tro}) \qquad \text{(III.1)}$$

where f is the transmitter frequency and c the velocity of light, and Δs_{tro} partially depends upon weather conditions in the vicinity of the station.

An initial expression for Δf_{tro} was derived in Ref. 1 on the basis of satellite--station geometry and the following simplifying assumptions about the troposphere:

(1) The refractivity N of air (where $N = 10^6(n - 1)$) is a continuous function of height above the earth but is independent of horizontal position and of time, within the region and the time interval of a satellite pass. This assumption is violated near a weather front but is otherwise fairly accurate.

(2) The N profile (height variation of N) can be approximated by a theoretical (quadratic) expression, decreasing from its value at the tracking station to zero at a specified height above the geoid, the "equivalent height" of the troposphere.

(3) Curvature of the signal path is small enough to be negligible, except close to the horizon where data are not used for geodesy.

The resulting expression for Δf_{tro} (Ref. 1) was incorporated into the orbit computing program, starting in January 1964. A value of Δf_{tro} is computed as a correction for the observed Doppler shift at each data point of each pass, using the geometry of a preliminary orbit and the theoretical refractivity profile described above. Initially a local seasonal mean value of the surface refractivity (Ref. 2) was used at each station as a starting point for the refractivity profile at the time of a pass; diurnal and weather variations were thus neglected. The equivalent height h_o of the troposphere was assumed to be 23 km at all stations.

The theoretical tropospheric contribution has the same sign as the Doppler shift itself throughout a satellite pass, but unlike the Doppler shift, its magnitude increases sharply at both ends of the pass, near the horizon. If no tropospheric correction is used, the Doppler residuals for any pass (observed minus theoretical Doppler shift) consistently show the sharp increase in magnitude near the horizon which theoretically characterizes the tropospheric effect.

This is illustrated by Figures III.1-1 and III.1-2, which show the Doppler residuals for two passes, one at medium elevation and one at a very low elevation, in each case both without and with the use of the tropospheric correction. Fig. III.1-1a and the upper graph in Fig. III.1-2 show the residuals when no tropospheric correction was made and also the theoretical tropospheric correction which should have been used for that pass. The residuals in both passes follow the shape of the theoretical tropospheric error curve (but displaced, since the frequency centering was affected by the large uncorrected refraction). Fig. III.1-1b and the lower curve in Fig. III.1-2 show the Doppler residuals which remained when the data were troposphere-corrected. The remaining errors are largely noise, indicating that the systematic errors in the two upper plots were tropospheric in origin.

Two further improvements have been made in computing the tropospheric effect.

First, the local refractivity of air at the tracking station is computed (Ref. 3) from the equation

$$N = \frac{77.6}{T} \left(P + \frac{4810\ e_s \times (RH)}{T}\right),$$

using weather data sent by the stations along with the Doppler data. This improvement went into effect in May 1964. In the above equation, T is temperature in degrees Kelvin, P is atmospheric pressure and e_s the saturation pressure of water vapor at the temperature T, both in millibars; RH is the relative humidity. The use of weather data provides the correct starting point for the refractivity profile and should result in a more accurate correction. This is especially important in warm, humid weather. The diurnal variation of surface refractivity in the Washington area is likely to be only about 10 N units (3 per cent) peak-to-peak in winter (winter weather effects also being small). In summer, diurnal variations are two or three times as large as this, while a cold front arriving in summer can drop the surface refractivity by 50 N units (15 per cent) within a few hours.

The second improvement stems from a recent study of refractivity profiles obtained from observed upper atmosphere data (Ref. 4). It was found that the increase in path length produced by the theoretical (quadratic) refractivity profile is in general not quite equal to the increase produced by an observed profile starting at the same surface refractivity. The ratio between the theoretical and observed effects is, however, a linear function of the surface refractivity for all the 34 profiles which were examined, regardless of geographic location, station altitude, or season. This relation can provide a simple correction factor for improving the quadratic profile results, and is now being introduced into the computation of the tropospheric correction to the Doppler shift.

The tropospheric effect, if uncorrected, produces an along-track error in position only if data are not symmetrical about the point of closest approach. It always produces an error in the apparent range from station to satellite at closest approach. Since tropospheric refraction steepens the slope of the observed Doppler shift vs. time curve, the uncorrected troposphere always makes the station appear closer to the orbit than it actually is. The amount of range error for a given state of the troposphere is a function of the maximum satellite elevation angle during the pass, and also depends on how much data near the horizon is included in the computation.

Figure III.1-3 shows this theoretical range error as a function of pass elevation, for data cut-off angles of 5°, 10° and 15° respectively, computed from the corrected quadratic profile as described above, and using a surface refractivity of 320, which is approximately an average value. The tropospheric range error for a 45° pass is 12 meters if all data are deleted below 15° elevation at both ends of the pass; but is more than twice as large (26 meters) if data are retained down to 5° elevation. The errors are considerably larger than this for lower elevation passes.

Values of surface refractivity which are encountered seldom differ by more than 20 or 25 per cent from the nominal value of 320 used here. A change of 20 per cent in the surface refractivity produces a theoretical change of not quite 10 per cent in the resulting range error. Thus most of the surface conditions which are actually encountered would result in theoretical errors within ± 10 per cent of those in Fig. III.1-3. Fig. III.1-4 shows the effect of temperature and humidity on refractivity, at an atmospheric pressure of 1000 millibars.

It is estimated that the latest form of the tropospheric correction can remove at least 85 or 90 per cent of the tropospheric effect, leaving range errors not more than 10 or 15 per cent of those shown in Fig. III.1-3. If the data are deleted below 10° satellite elevation, the maximum error is thus less than 5 meters, and the RMS error is probably around 3 meters when taken over all passes.

Residual uncorrected errors have several possible sources. The curvature of the signal path has been neglected; but this is extremely small at angles above 5°. Probably more important is the fact that the theoretical and the observed mean profiles are not quite the same shape. As the elevation of the signal path is lowered from 90° toward the horizon, the lower layers of each profile are stretched more than the upper layers, and this differential stretching must have slightly different effects on the theoretical and the actual profiles.

No account has been taken of the difference of an actual instantaneous refractivity profile from the mean profile at that location (e.g. in the presence of a weather front), or of the changes which may occur during a pass. These factors will be sources of noise in the data, but should not have a biasing effect when a large number of satellite passes is being considered.

Further study of the refractivity profile, especially of the lower part of it, is very desirable if the expression is to be refined.

Refraction errors are biased in their effect upon the slant range, but are random in their effect upon position when considered for all passes.

References - Section III.1

1. Hopfield, H. S., The Effect of Tropospheric Refraction on the Doppler Shift of a Satellite Signal, J. Geophys. Res., 68, pp. 5157-5168, September 15, 1963.

2. Bean, B. R. and J. D. Horn, On the Climatology of the Surface Values of Radio Refractivity of the Earth's Atmosphere, National Bureau of Standards, Report 5559 (Boulder Laboratories), 3 March 1958.

3. Smith, E. K. and S. Weintraub, The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies, Proc. IRE 41, 1035-1037, 1953.

4. Hopfield, H. S., Improvement of the Tropospheric Correction for Doppler Data, Based on a Study of Upper Air Meteorological Data, Applied Physics Lab. Report TG-646, January 1965.

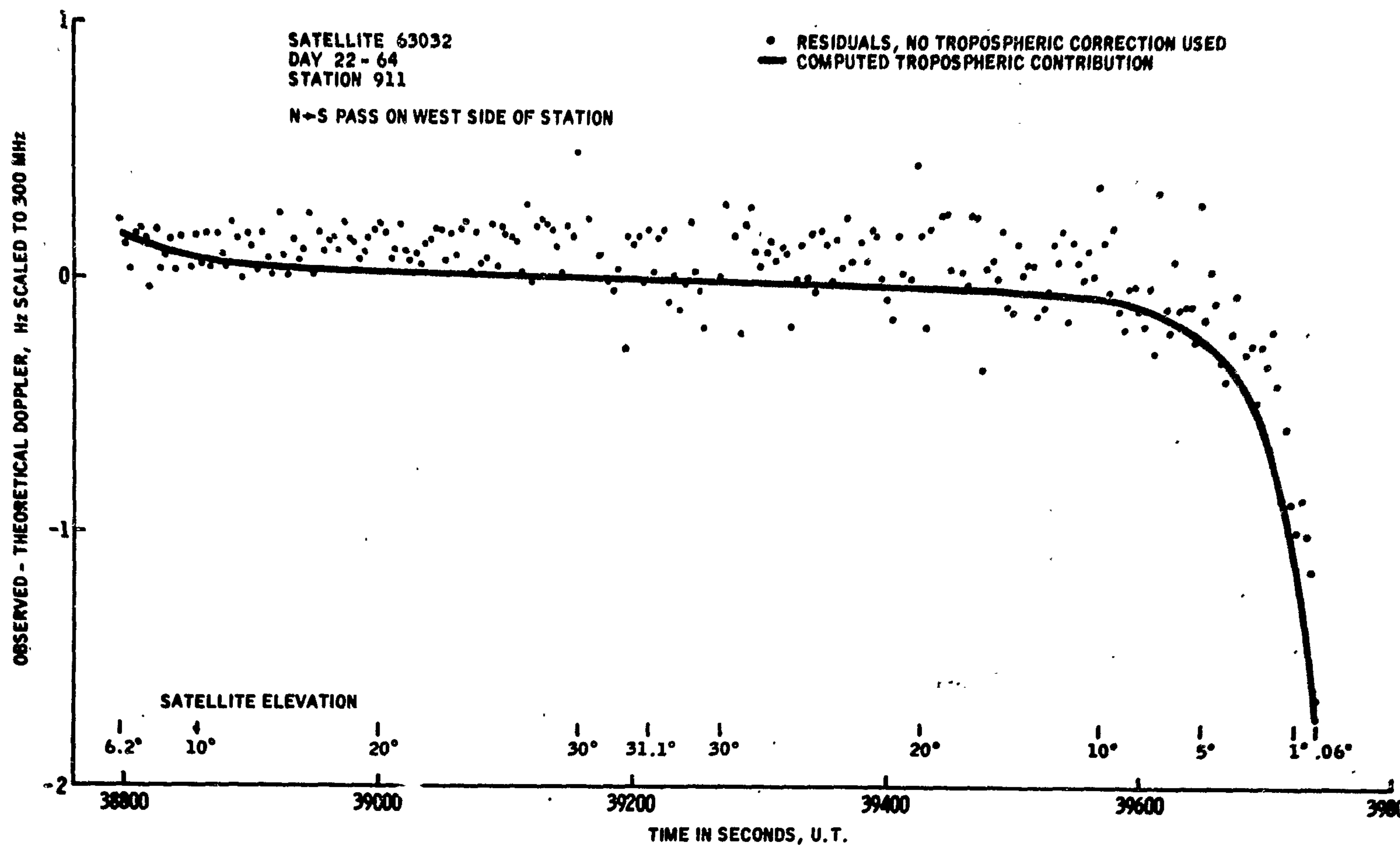

Fig. III. 1-1a DOPPLER RESIDUALS FOR A PASS WITH 31° MAXIMUM ELEVATION, WITHOUT THE USE OF TROPOSPHERIC REFRACTION CORRECTION.

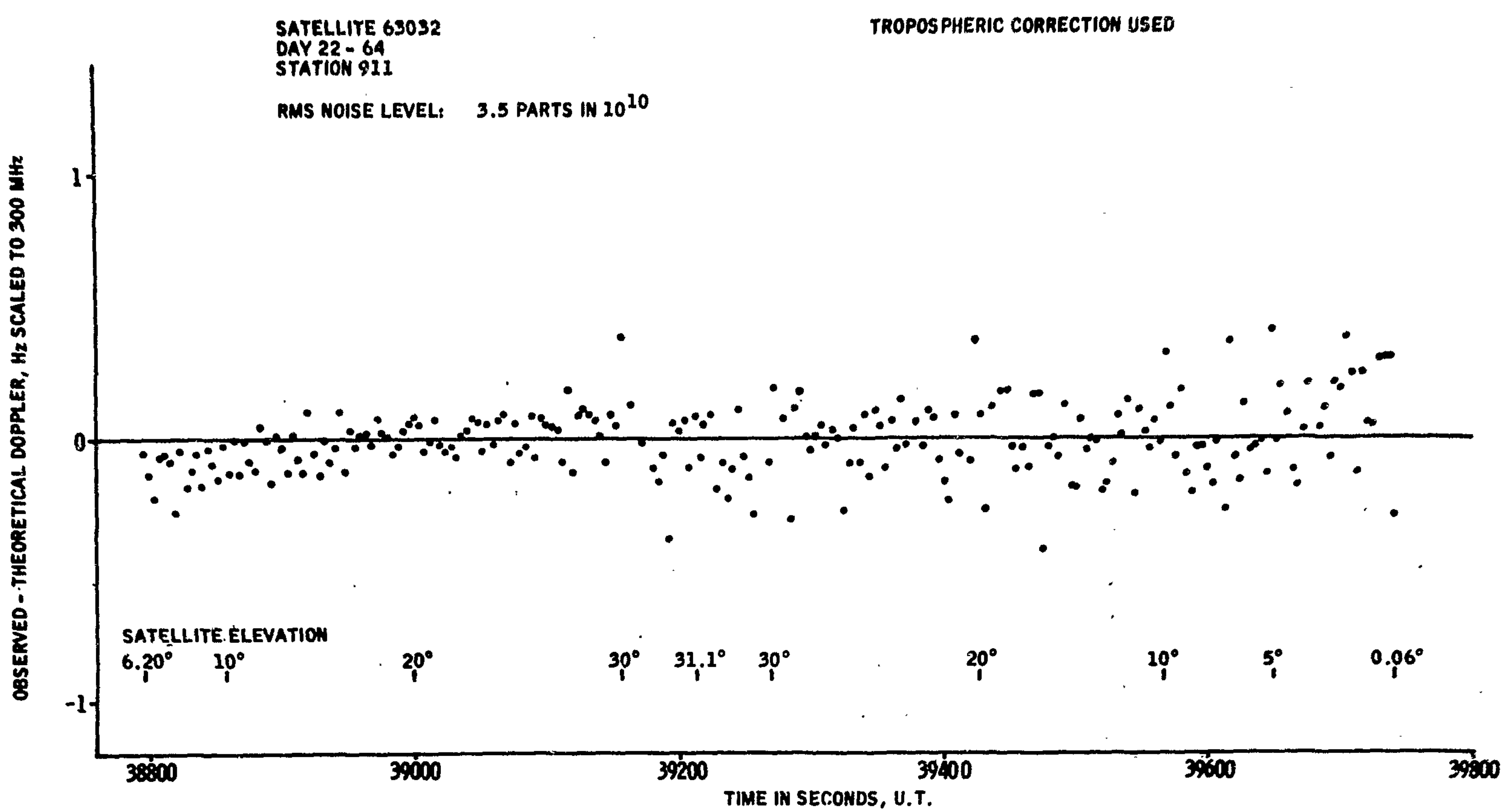

Fig. III. 1-1b DOPPLER RESIDUALS FOR A PASS WITH 31° MAXIMUM ELEVATION, WITH THE USE OF TROPOSPHERIC REFRACTION CORRECTION.

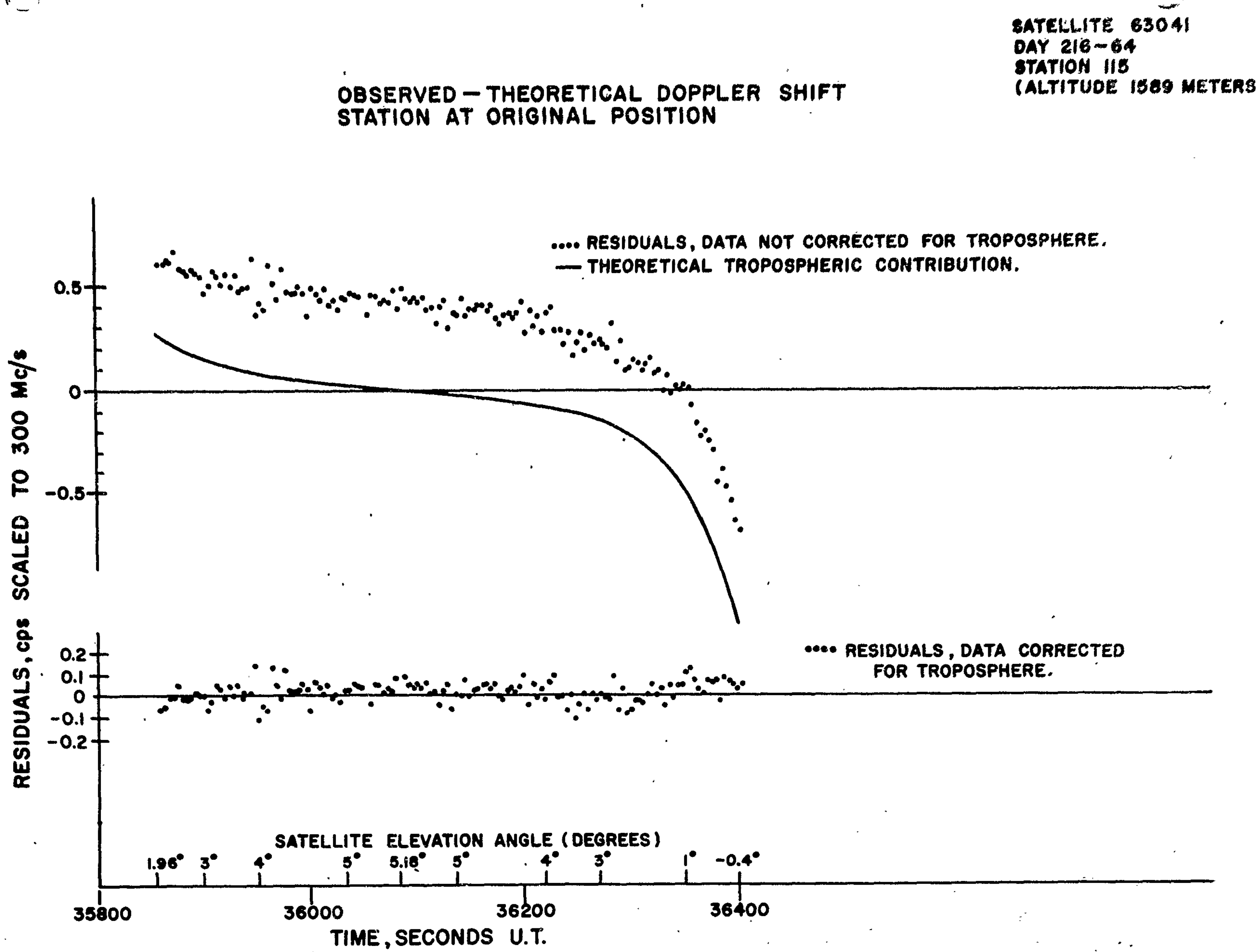

Fig. III. 1-2 DOPPLER RESIDUALS FOR A PASS WITH 5° .18 MAXIMUM ELEVATION, (a) WITHOUT AND (b) WITH TROPOSPHERIC REFRACTION CORRECTION.

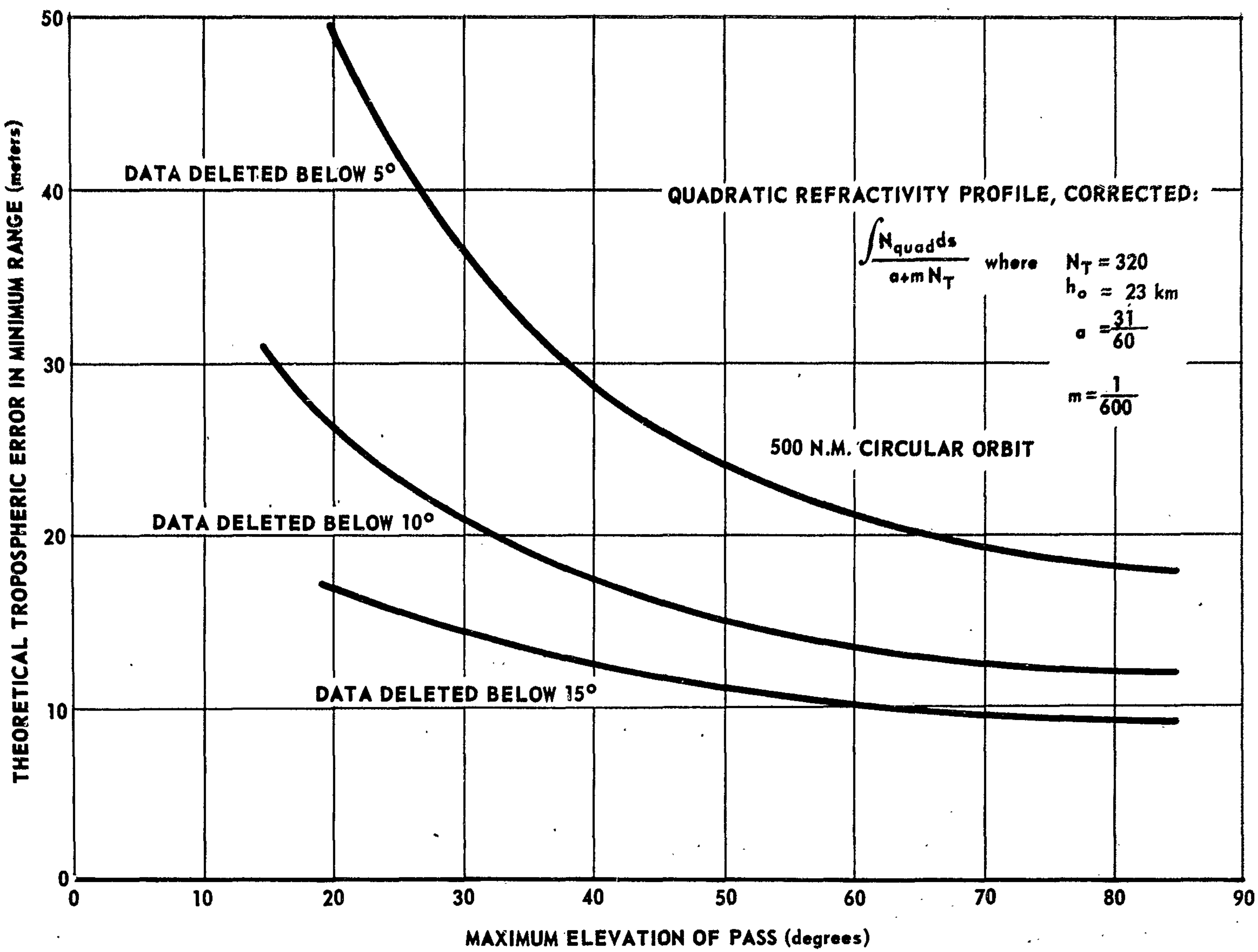

Fig. III. 1-3 RANGE ERROR FROM UNCORRECTED TROPOSPHERIC REFRACTION FOR DIFFERENT DATA CUT-OFF ANGLES.

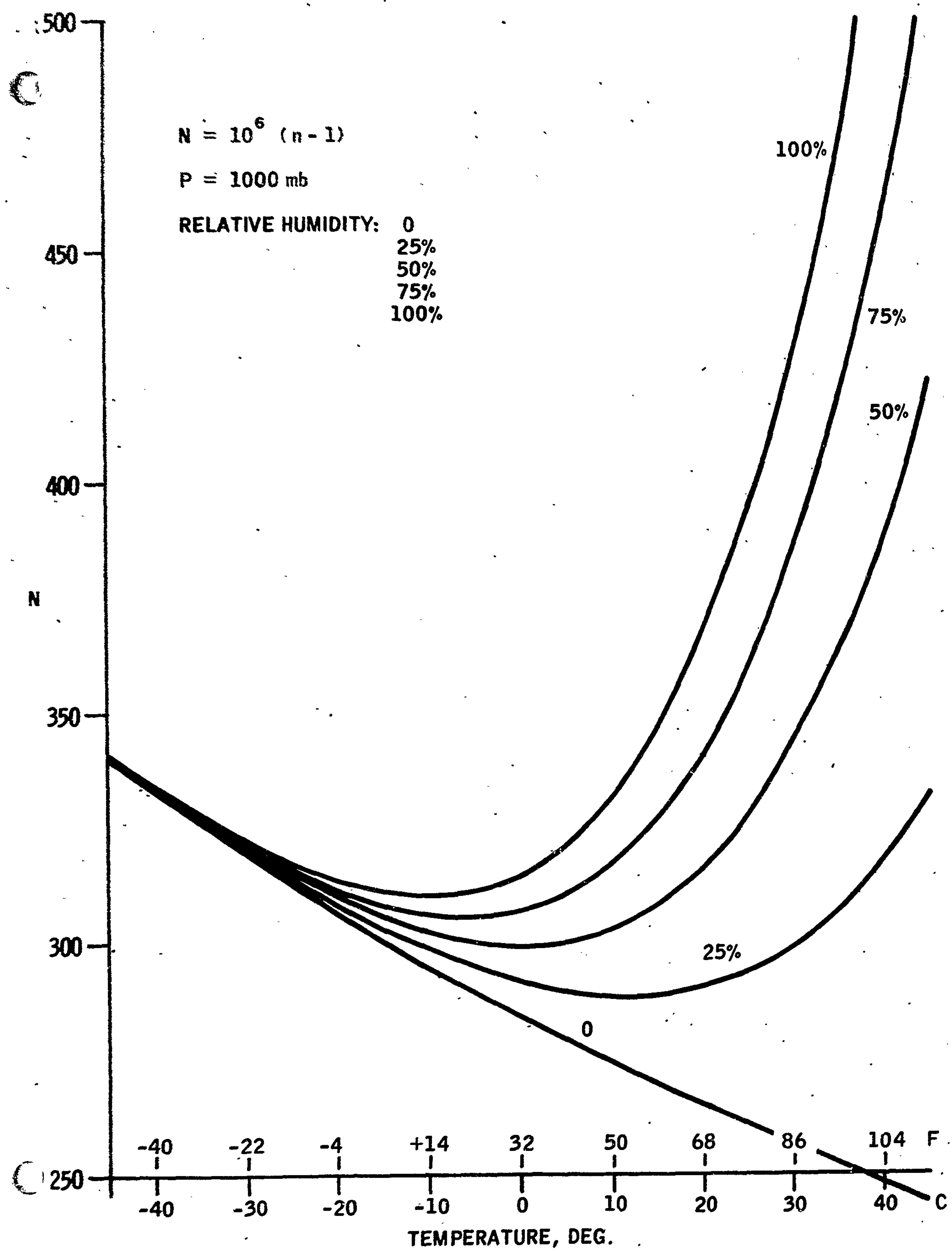

Fig. III. 1-4 TROPOSPHERIC REFRACTIVITY.

III.2. Ionosphere Contributions to the Errors.

Although the analogy with optical refraction for the ionosphere is less direct than with tropospheric refraction, treatments of the ionosphere usually are cast within this analogy. In considering the ionosphere, this requires:

(1) that the index of refraction be less than unity,

(2) that equivalent medium be dispersive (the index is frequency dependent),

(3) that the equivalent medium be optically active (a different value for the index of refraction for different polarizations of the electromagnetic signal).

The analogy with optics and the resulting expressions for the index of refraction and contributions to the Doppler shift have been extensively studied[Refs. 1-4] and the results are briefly summarized below.

For VHF frequencies and higher, at least up to several kilo-megacycles, the index of refraction and the contributions to the Doppler shift after correction for the troposphere can be expanded in a power series. For the Doppler shift, this correction has the form:

$$\Delta f = \Delta f_v + \frac{a_1}{f} + \frac{a_2^{(\pm)}}{f^2} + \frac{a_3}{f^3} + \text{higher orders}$$

where

Δf_v = Doppler shift which would occur in a vacuum,

a_1/f = first order contribution of ionosphere which is proportional to the time derivative of the electron density integrated along the slant range vector from station to satellite,

$a_2^{(\pm)}/f^2$ = second order contribution or Faraday term which depends upon whether the polarization is right circular (±) or left circular (-) and is proportional to the time derivative of the integral of the electron density and the component of the earth's magnetic field along the direction of propagation,

a_3/f^3 = third order contribution which depends upon various powers of the electron density and its spatial gradient.

Considering the Doppler shift at any single frequency, the first order contribution is by far the largest. It is always eliminated in data used for geodetic purposes by combining measurements taken simultaneously on two coherent frequencies. It appears that the second order term is considerably smaller than the third order term and can be considered negligible. The third order term can be non-negligible when:

(1) the electron density is very large;

(2) the gradient of the density is large.

Clearly, if the density is large, its gradient is most likely large. However, even when the density is small, its gradient can be large at times when there are ionospheric disturbances.

The third order term is usually negligible and currently its contribution is neglected for all passes. However, considerable study is underway to understand its character more fully. The principal reason is that in a few years the solar activity is expected to become much higher (approaching the sunspot cycle maximum while currently we are near the sunspot minimum). To examine its character more fully, data taken in previous years are being studied and some current results are summarized below.

In order to determine the effect of the residual ionospheric errors on station position a series of calculations was run. The following list summarizes the essential steps of the calculation. Steps (1) and (2) were carried out at the Defense Research Laboratories.(Ref. 4)

(1) Single frequency doppler data at 54, 150, 324, and 400 mc. were obtained from Station 092 (operated by DRL) at Austin, Texas.

(2) Assuming a_2 to be zero, a system of equations (of the form of the power series truncated at a_3) in three unknowns (Δf_v, a_1, and a_3) were solved using three of the four measured dopplers. In this way values of a_3 referenced to 54 mc. were obtained.

(3) The values of a_3 were converted to the residual error that would have been present in the doppler data from the first order refraction corrected 150-400 mc. pair of frequencies.

(4) Using the geometry of the actual tracking station on a rotating earth, and a circular orbit which closely approximated the actual orbit during the observed pass, the change in station position which would result from the observed set of Doppler errors for the pass was determined.

Figures III.2-1 through III.2-5 illustrate the time behavior of a_3 during a pass for several passes, and also summarize the position errors introduced by a_3. Table III.2-1 summarizes the essential results of the calculation. The cut-off angle means the lowest elevation angle of the satellites used in the calculations.

The following points are noteworthy:

(1) the a_3 data were accumulated during the months of March and April of 1962 and were mainly from afternoon passes.

(2) From the data analyzed it was observed that residual ionosphere refraction error could produce up to 20 meters total error.

(3) The 20 meter total error was much larger than the average value and was found to be correlated with an ionospheric disturbance (sudden enhancement of the electron density) just prior to the pass.

Omitting the 20 meter value because of its correlation with a known disturbance, Table III.2-2 gives a summary of the dependence of the position error upon frequency pair used and upon the cut-off angles. For 10° cut-off, which is normally used, the residual error is 2.6 meters with the 150-400 megacycle pair; the statistical uncertainty in this estimate is 1.4 meters.

During the period when the a_3 measurements were made, the solar activity was down considerably from its value at sunspot maximum. The intensity S of solar radiation in the 10.7 cm. band is a good measure of solar activity; in units of 10^{-22} watts/m^2 cps, S had an average value of about 230 in 1958 and of about 70 in 1964. During the period applying to Table III.2-2, the average was 98. The value of S is expected to have a profound effect upon ionospheric refraction.

An analysis of ionospheric data from 1953 through 1959 (previous minimum to maximum of S) indicates that the electron density may increase by as much as a factor of 4 at the next sunspot maximum. The effects due to a_3, on the basis of current theory, go about as the square of the electron density. Hence, our present best estimate is that the values in Table III.2-2 should be multiplied by 16 to give position errors during maximum solar activity. The next maximum is expected to occur around 1968 or 1969.

If we apply a factor of 16 to the values in Table III.2-2, we get position errors in the range of 20 to 60 meters when using a 10° cut-off with the 150-400 mc. pair. However, this number does not represent the situation adequately, for the following reasons:

1. The factor of 16, deduced from ionosonde data, reflects only the expected change in the maximum electron density in the ionosphere. By contrast, refraction effects upon Doppler frequencies depend partly upon the total integral of the electron density, and may vary even more rapidly with solar activity; that is, the correct factor may be even greater than 16.

2. The data have been collected from only one site, which has a latitude of about 30°. Roughly, we may expect smaller errors at higher latitudes and vice versa. Experiments to obtain data similar to those in Figs. III.2-1 at other latitudes are expected to give some results in 1965.

3. Nighttime passes are expected to give smaller errors on the average than those used in preparing Table III.2-2, by a factor that may be as much as 10. Thus nighttime passes are expected to have acceptably small errors even at maximum solar activity. We do not believe that much weight should be given to this fact, however, because we believe that it is important to have data distributed uniformly in time.

In summary, using dual-frequency Doppler data with the 150-400 mc. pair, we may expect many passes to contain high-order refraction errors in the range of several tens of meters. Without corrective action, it will probably be necessary to decrease either geographic coverage or time coverage during maximum solar activity, for say three years around 1969-1971. We can think of three types of possible corrective action, all of which are being investigated. These are:

1. We can use three frequencies in order to correct the a_3 term as well as the a_1 term presently being corrected. This approach does not seem desirable because it would complicate the ground equipment considerably.

2. We can use higher frequencies; use of the 324-972 pair, which is technically feasible, gives errors due to a_3 of only a few meters at solar maximum. This would not increase the complexity of a ground station electronic system, although it would require replacement of some components.

3. We can calculate the a_3 term from the a_1 term which is currently being measured. This is the most desirable possibility if it works, since it does not require any change in station equipment. However, we do not know how accurate this calculation will be; studies of this calculation are being made. If it is only accurate to 25% of a_3, which is quite possible, it is not accurate enough.

References - Section III.2

1. Stratton, J. A., Electromagnetic Theory, McGraw-Hill Book Company, New York, 1941.

2. Guier, W. H., Ionospheric Contributions to the Doppler Shift at VHF From Near-Earth Satellites, Applied Physics Lab. Report CM-1040, July 1963.

3. Pisacane, V. L. and J. D. Stuart, The Effect of Collision on the Group Refractive Index of the Ionosphere, Applied Physics Lab. Report TG-649, January 1965.

4. Defense Research Laboratory Ionospheric Studies, DRL Report 487, J. F. Willman, August 1962, DRL Report 491, J. F. Willman and J. F. Doyle, February 1963.

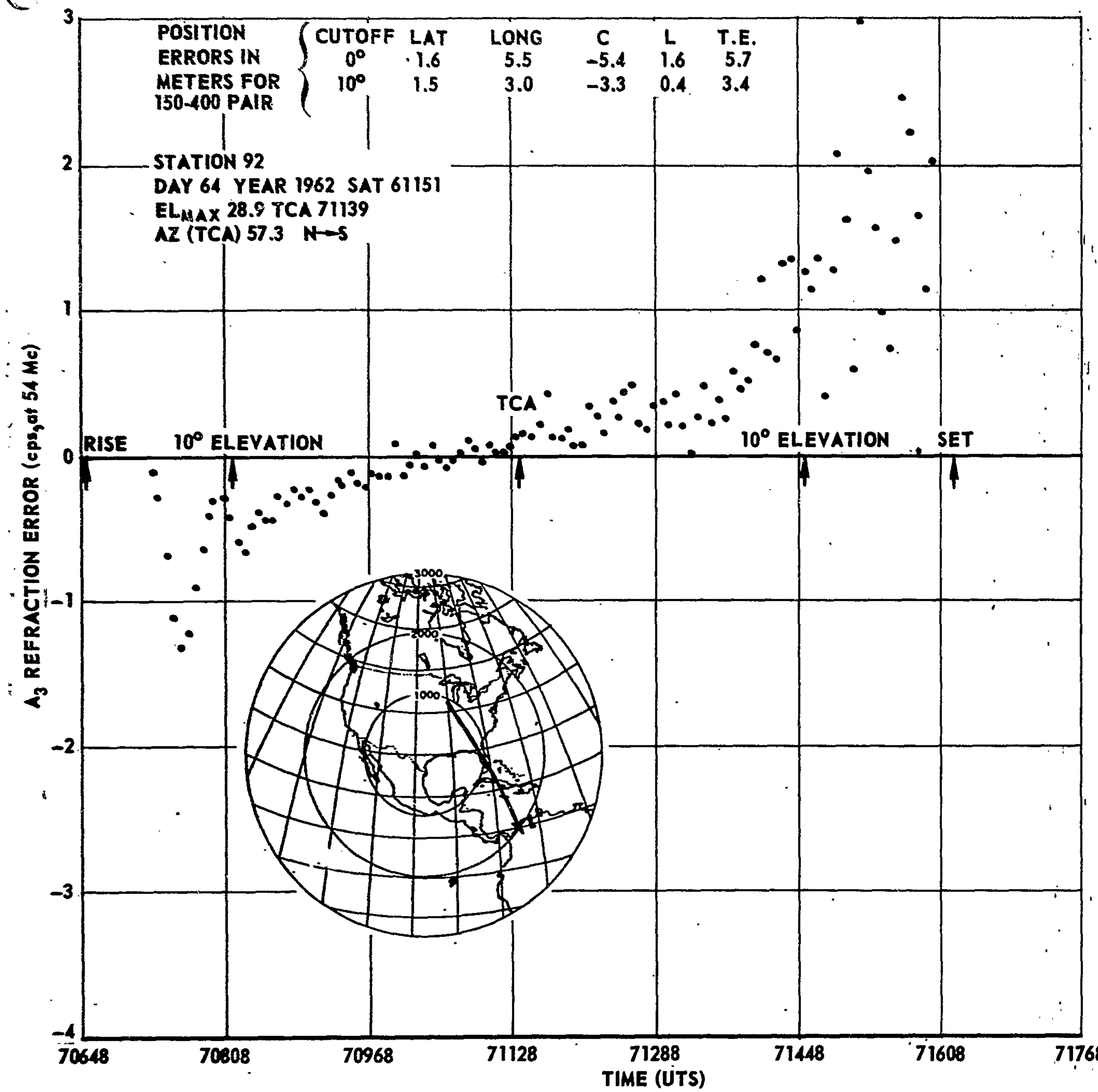

Fig. III. 2-1 THIRD-ORDER IONOSPHERIC REFRACTION AND THE CORRESPONDING STATION POSITION ERRORS FOR A PASS HAVING THE GEOMETRY SHOWN IN THE INSERT.

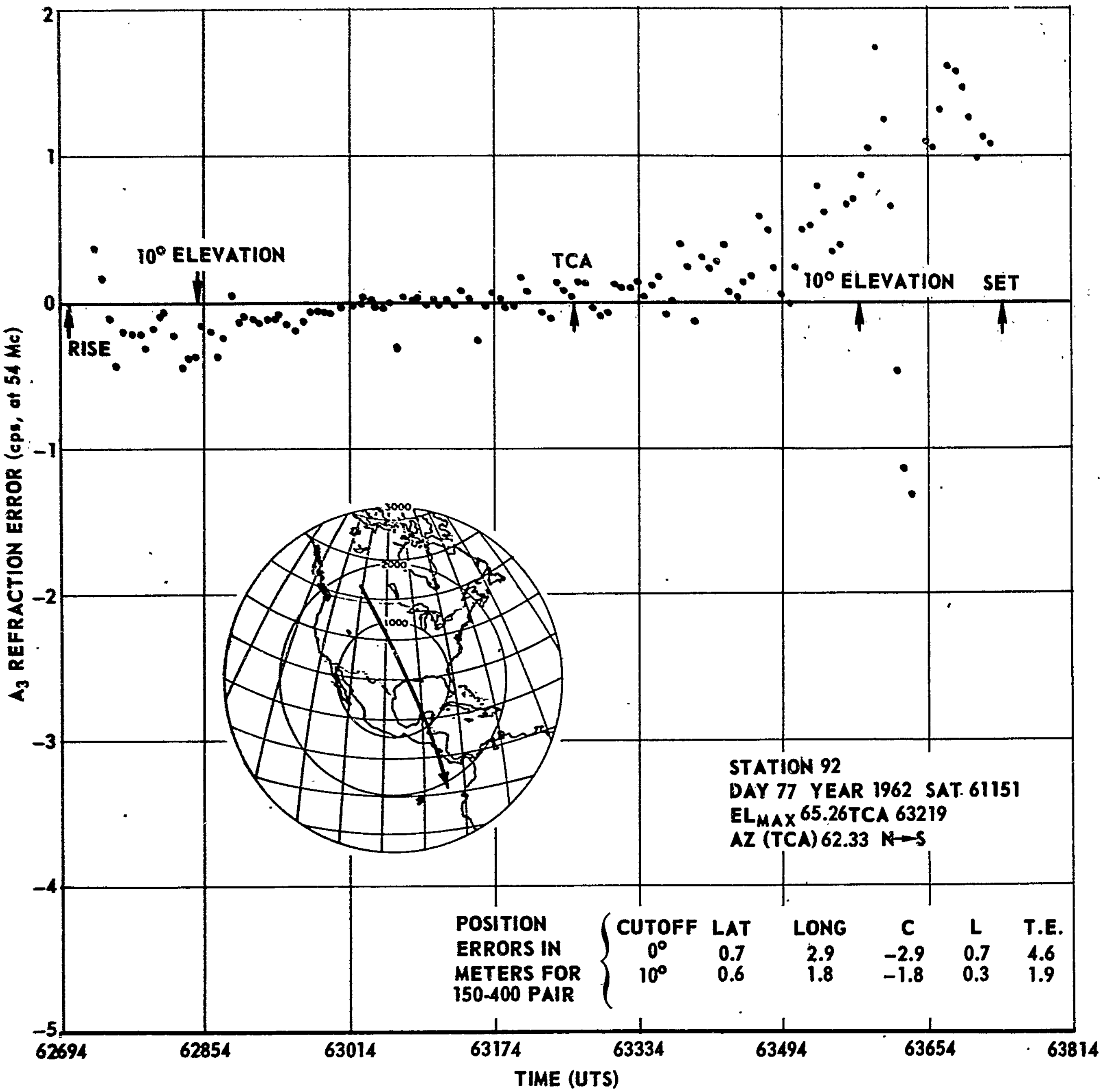

Fig. III.2-2. THIRD-ORDER IONOSPHERIC REFRACTION AND THE CORRESPONDING STATION POSITION ERRORS FOR A PASS HAVING THE GEOMETRY SHOWN IN THE INSERT.

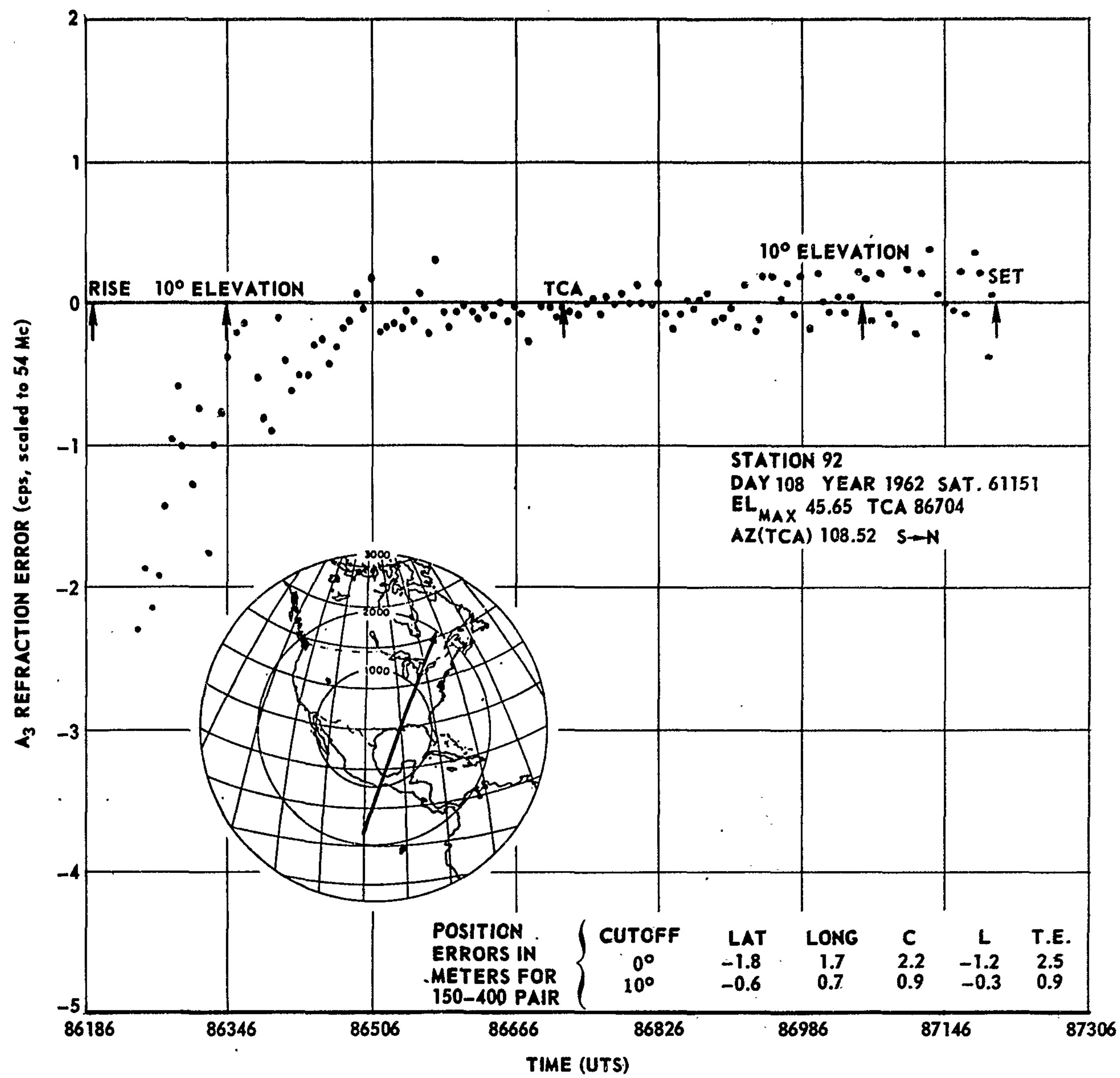

Fig. III.2–3. THIRD–ORDER IONOSPHERIC REFRACTION AND THE CORRESPONDING STATION POSITION ERRORS FOR A PASS HAVING THE GEOMETRY SHOWN IN THE INSERT.

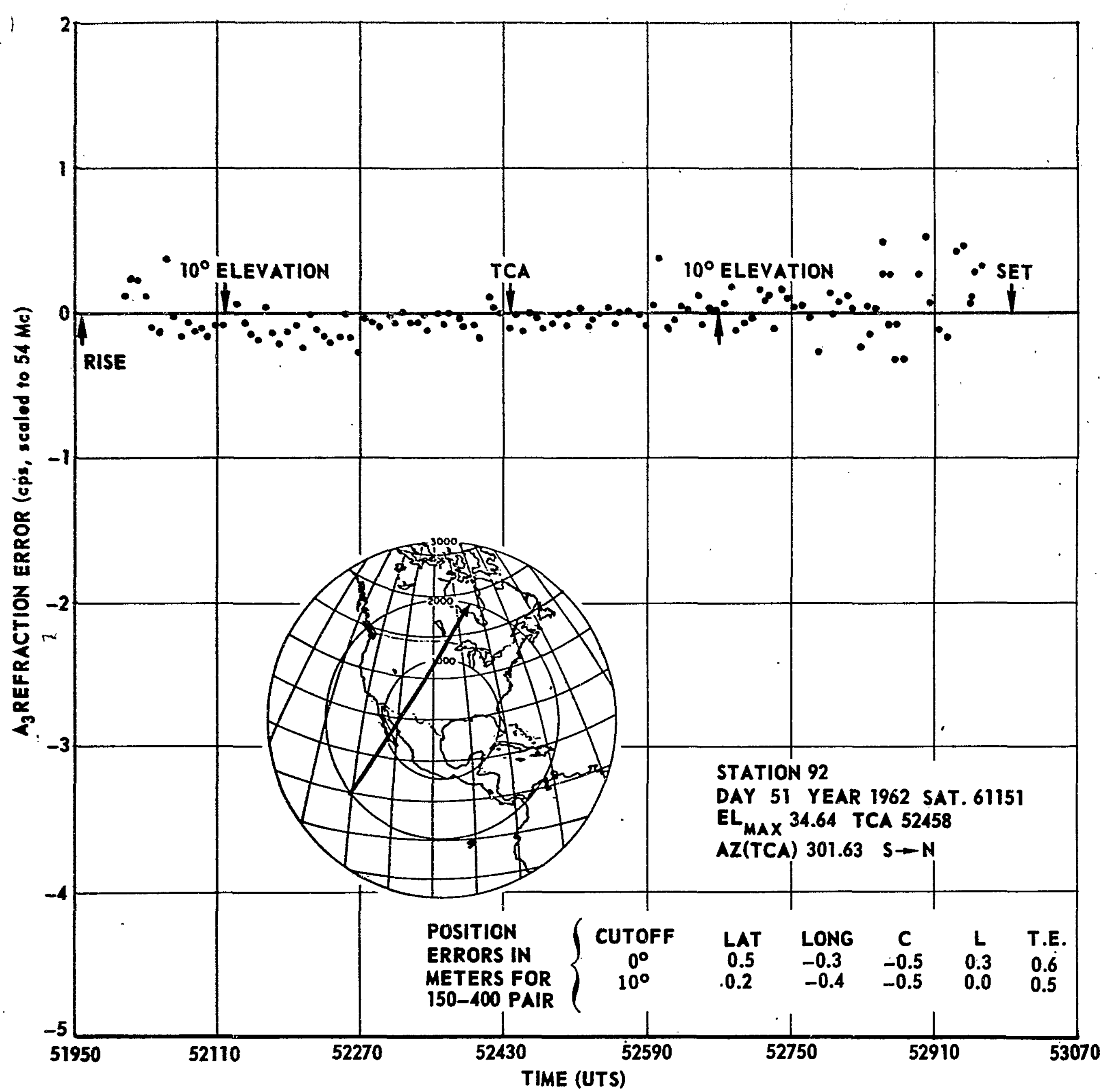

Fig. III.2–4. THIRD–ORDER IONOSPHERIC REFRACTION AND THE CORRESPONDING STATION POSITION ERRORS FOR A PASS HAVING THE GEOMETRY SHOWN IN THE INSERT.

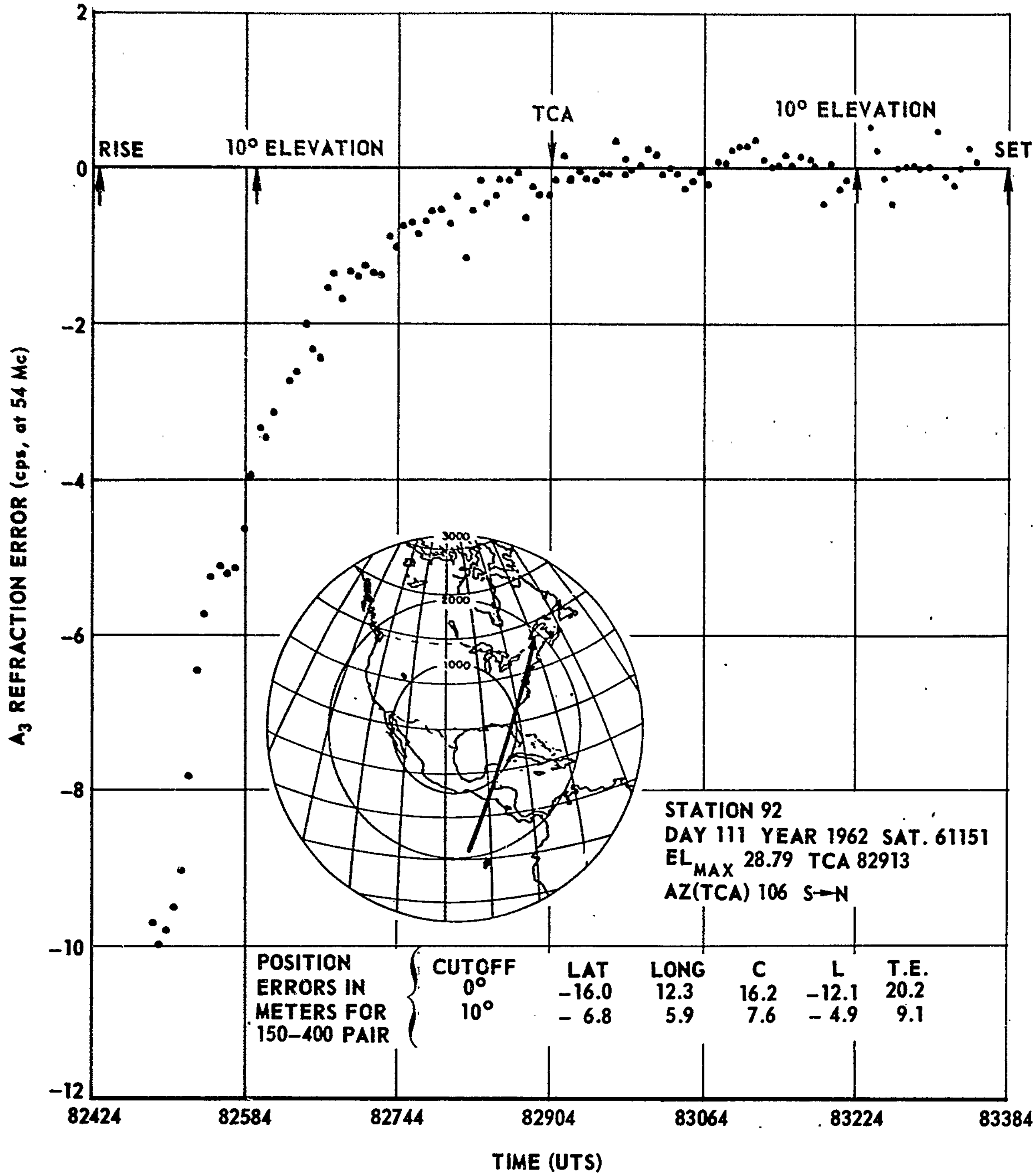

Fig. III.2–5. THIRD-ORDER IONOSPHERIC REFRACTION AND THE CORRESPONDING STATION POSITION ERRORS FOR A PASS HAVING THE GEOMETRY SHOWN IN THE INSERT

Table III.2-1

Position Errors Due to High Order Ionosphere Refraction

For Typical Daytime Passes

Year 1962
Months March-April
Station DRL - Austin, Texas

Local Time	Maximum Satellite Elevation	Azimuth at Closest Approach	Direction of Satellite Motion	Position Error, Meters 0° Cut-off	Position Error, Meters 10° Cut-off
$8^h\ 34^m$	34.6	301.6	S → N	0.6	0.5
$14^h\ 20^m$	22.3	105.6	S → N	8.5	3.6
$16^h\ 07^m$	45.0	298.3	S N	5.1	2.9
$16^h\ 12^m$	31.9	107.2	S → N	11.8	5.5
$17^h\ 01^m$	28.8	106.0	S → N	20.2	9.0
$17^h\ 50^m$	26.0	105.9	S → N	6.9	3.2
$18^h\ 05^m$	45.7	108.5	S → N	2.5	0.9
$18^h\ 34^m$	60.4	298.1	S → N	3.7	2.0
$11^h\ 04^m$	23.2	55.7	N → S	2.7	1.1
$11^h\ 33^m$	65.3	62.3	N → S	3.0	1.9
$13^h\ 45^m$	28.9	57.3	N → S	5.7	3.4
$15^h\ 32^m$	36.5	252.1	N → S	5.9	4.2
$15^h\ 57^m$	13.1	50.6	N → S	1.1	1.0
$16^h\ 35^m$	23.0	253.8	N → S	7.0	3.4

	0° Cut-off	10° Cut-off
Average Error:	5.0	2.6
Standard Deviation: (in estimate of average error)	3.0	1.4

Table III.2-2

Estimated High Order Refraction Error as a Function of Frequency Pair

Frequency Pair	0° Cut-off Average Error in Meters*	10° Cut-off Average Error in Meters*
150-400	5 ± 3	2.6 ± 1.4
324-648	0.4 ± 0.3	0.2 ± 0.1
324-972	0.2 ± 0.1	0.1 ± 0.1
648-972	< 0.1	< 0.05

*The uncertainty results from the statistics of the data sample used, and does not reflect possible inadequacies in the theory of high-order refraction.

IV. LIMITATIONS UPON SYSTEM ACCURACY IMPOSED BY SYSTEM COVERAGE

By the system coverage, we mean the number and distribution of the satellite orbits for which data are available and of the ground stations from which measurements are made. For dynamical geodesy, coverage limits the number of gravity components that can be estimated and hence limits the system accuracy. For geometrical geodesy, accurate orbits are not needed except over long distances (thousands of kilometers), and coverage in this sense is not so important. Hence the discussion of this section deals only with dynamical geodesy.

IV.1. Ground System Coverage

The aspect of dynamical geodesy which imposes the most severe requirements upon ground system coverage is determining the coefficients of the non-zonal gravity harmonics. These coefficients may produce more than one type of effect, but one effect that they always produce is an orbital perturbation whose period is close to (24/m) hours, where m is the order of the harmonic.*

---

*A non-zonal harmonic of degree n and order m has the analytic form $[Y_{n,m}(\theta)/r^{n+1}]\cos m(\lambda-\lambda_{n,m})$. In this, r denotes radius from the center of the earth, and θ and λ denote latitude and longitude. $\lambda_{n,m}$ is the longitude of the symmetry plane of the harmonic. $Y_{n,m}$ is the associated Legendre function; its form concerns us here only to the extent of noting that is is a function only of latitude and hence that it does not vary with time(at a particular point in space) as the earth rotates. The trigonometric factor containing the longitude is a periodic function of the time when the rotation of the earth is considered; the period obviously depends only upon the order m.

Ref. 1 estimated that the analytical error will be reduced to 10 meters, the estimated current level of observational error, when all harmonics through degree 16 have been inferred on the basis of adequate orbital coverage. This means that we will be dealing with perturbations having periods as short as (24/16) = 1.5 hours.

Ref. 2 dealt with the question of station deployment needed in order to analyze for all harmonics through degree 8, for which the perturbation period is 3 hours. The criterion adopted in that reference was that the station configuration should provide tracking data with a maximum data gap of no more than 1 hour, for satellites of any inclination with an altitude of 600 n.m. This insured at least three observations, and usually more, within the period of any perturbation which the system was designed to study. This density of observations should be adequate to allow accurate resolution of the perturbation periods for periods of 3 hours or longer.

On the basis of that study, it was recommended that the network of stations intended for gravity research be extended from the present 12 to 14, the additional stations to be in Antartica and the Canary Islands. With this extended network, the criterion of data coverage would be satisfied with negligible exceptions.

If the analysis is extended from harmonics of the eighth order through harmonics of the sixteenth order, it might be expected that we would have to extend the coverage in order to halve the interval between observations. This would be so if we did not have more powerful means of analysis than we did when Ref. 2 was prepared. At that time, the longest time span used for a single arc was 24 hours, at least for purposes of analyzing the non-zonal harmonics. Now we know that we can hold arcs of least 72 hours (3 days) with the computing accuracy needed for the non-zonal analysis. Thus, in examining data density, we can consider what happens if all of the data from three days are plotted on the same 24-hour time base, making of course the correct adjustment to the data from the later days. Unless the orbital period is very close to the period of the perturbation or to a multiple of it*, corresponding observations on different days will lie at different phases of the pertubation, and we have the same result that we would get from more dense coverage using 24-hour arcs.

We do not have to worry about the possibility that the orbital and perturbation periods are close to each other, keeping us from seeing different phases of the perturbation. If the periods are close, we have a resonant condition (Ref. 3) which gives a perturbation with a characteristic long-period beat. Also the resonance greatly magnifies the size of the perturbation produced on the orbit, easing the data requirements needed to study the perturbation. We have been able to study resonances for orders 13 and 14 with data from the present system. The limitation in studying harmonics of these orders has been lack of satellites experiencing strong resonance, not low density of data for the satellites that do experience them.**

---

*We do not need to consider the case involving a multiple until the analysis is extended to harmonics of order around 25.

**There is one possible problem with near-resonance. Some types of resonance produce an oscillation with orbital period, which is amplitude modulated at the beat period. These may be difficult to analyze with the present data density, but at present we hope that this problem is small enough not to require extension of the ground system.

In summary, to the best of our current knowledge, the present ground system extended by stations in Antarctica and in the vicinity of the Canary Islands will be adequate for gravity analysis through harmonics of degree and order 16. At this level, the analytic error is expected to be about 10 meters. Extension of the analysis to this point will require additional satellites in suitable chosen orbits.

References - Section IV.1

1. Guier, W. H. and R. R. Newton, Status of the Geodetic Analysis Program at the Applied Physics Laboratory, Applied Physics Lab. Report TG-652, January 1965.

2. Letter TS-612, Director, Applied Physics Laboratory to Chief, Bureau of Naval Weapons, Ground Station Deployment in Support of TRANET, 4 February 1964.

3. Yionoulis, S. M., Resonant Geodesy, Applied Physics Lab. Report TG-633, December 1964.

IV.2. Configuration of Orbits

Limitation of system accuracy due to the ground station configuration arises from the need to have data with sufficiently uniform and dense time distribution. Limitation of accuracy due to orbital configuration arises from the fact that changes in different coefficients in the geopotential can have almost identical effects upon the inferred motion of a satellite in a particular orbit. Such coupled coefficients can be separated accurately only by using orbits with sufficiently different parameters.

Of the six parameters of an orbit, one is a time parameter and enters only into determining the phase of certain perturbations. Of the five geometric parameters, the argument of perigee and the longitude of the node change secularly*; these affect the period of many perturbations but do not affect their size except in minor ways. Only the semi-major axis, the eccentricity, and the inclination have a major effect upon the magnitude of a perturbation.

Up to the present, all satellites used for the geodetic analysis of Doppler data have had almost zero eccentricity and altitudes near 1000 km.; in fact, this is true of almost all satellites used for geodetic analysis with any type of data. The only significant variety in orbits has come from the inclination. We consider first the extent of analysis possible with orbits differing only in inclination.

---

*For certain orbits that have not yet been encountered, it may be that the argument of perigee has an oscillatory motion. The answer to this question is not known rigorously.

This question was considered in Ref. 1. The obvious factors are the following:

1. The characteristic period of the perturbation produced by a harmonic in the geopotential is (24/m) hours. For m = 0, the period is infinite as implied (a secular perturbation) when n is even. For m = 0 and n odd, high order effects keep the period from being infinite, but it is of the order of thousands of satellite revolutions.

2. For a given m, the effects of coefficients with even n are qualitatively similar. Likewise, the effects of coefficients with odd n are qualitatively similar. Coefficients with different parity of n are qualitatively dissimilar. (This rule holds only for n values fairly close together.)

It is easy to separate perturbations of different period, provided that the data density is adequate. Likewise, it is easy to separate the effects of coefficients with different parity of n. However, coefficients with the same m and the same parity of n produce perturbations which are quite similar for a given inclination but whose magnitudes change with inclination in different ways. Thus, if we have orbits with k different inclinations, we can infer the harmonics for k different even values of n and also for k different odd values of n, all for each value of m. Naturally we start with the smallest values of n and work up to values as large as this rule permits.

Since different coefficients with the same m and the same parity of n do not produce identical, but only similar, effects it is conveivable that we could infer more harmonics than this rule allows. If we should try, we would find that harmonics with the same m and parity of n have high correlation; this is a complicated way of saying that objects that look alike are hard to tell apart. With error-free data ideally distributed, we could still get reliable results. With actual data, such results are suspect.

For example, with three inclinations, our rule allows inferring the coefficients marked with x in the following scheme:

n \ m	0	1	2	3	4	...
2	x		x			
3	x	x	x	x		
4	x	x	x	x	x	
5	x	x	x	x	x	...
6	x	x	x	x	x	...
7	x	x	x	x	x	...
8		x		x	x	...
9					x	...
...						...

This scheme makes no reference to the harmonic (0,0); this harmonic is proportional to the total mass of the earth, it is a special case that does not fit into these considerations, and it will be discussed later. The coefficients with indices (2,1) are zero in the ideal coordinate system for which the polar axis of the coordinate system coincides with the spin axis of the earth. These coefficients are known to be negligible compared with coefficients that we are currently trying to find, thus we are entitled to ignore (2,1) and to go through n = 8 for m = 1 if we wish. We prefer to infer values of the (2,1) coefficients as a check on the validity of the results, and to omit the last value of n for m = 1.

All coefficients with n = 1 are zero in the coordinate system used, in which the origin is at the center of mass of the earth by definition, and we have omitted these coefficients in the above scheme. However, the analysis requires that we revise the coordinates of the tracking stations as the gravity analysis is improved. Certain changes in station coordinates are equivalent to non-zero values of the n = 1 coefficients; thus we are required to consider changes that resemble coefficients with n = 1, and may have to omit the highest odd values of n for m = 0 and 1. In practice we have found it safe to use this n value for m = 0, that is, to deduce zonal harmonics through the seventh with 3 satellites, through the ninth with 4, and so on. In one test, we found it unsafe to do the same for m = 1; it is not known whether it is always unsafe or whether this was an accident of the particular data available.

To sum up the preceding paragraph, we have found it best to stop the m = 1 column at n = 6 instead of at n = 8, as the scheme shows it, for 3 satellites, at n = 8 instead of 10 for 4 satellites, and so on.

In practice, we have stopped at this value of n for all m (except m = 0), adopting instead the simple rule n = 2k where k is still the number of significantly different inclinations. Formally, this practice is required by the computing programs used. We could remove this program limitation by re-programming, but so far we have not for the following reason: The additional coefficients that we could find by lifting the restriction have high values of m. For high m and hence short period perturbations, the data requirements are high. The "data" are certain outputs from a body of highly accurate orbit determinations, and so far we have not had the time to do enough such determinations to support analysis for large m. The raw Doppler data needed do exist and will be reduced as time permits.

Aside from this problem, there are two reasons for departing from the simple scheme. One reason tends to decrease the number of coefficients that can be inferred, the other tends to increase it.

The first reason is simple to state. As the number of orbits and the number of coefficients inferred increases, it will not be possible indefinitely to avoid large correlations between all coefficients lying within the rule. With a limited number of inclinations, it is inevitable that two coefficients with the same m and parity of n will have almost proportional coefficients for two of the inclinations, so that only one of these can be found.

The cause that allows us to find more coefficients than the simple rule allows is the phenomenon of resonance. If an orbital period is nearly equal to the period of the perturbation caused by a particular value of m, a typical beat oscillation occurs with a large period and correspondingly large amplitude.

The periods of the satellites used so far in the Doppler analysis lie between (24/13) and (24/14) hours, thus the resonances encountered so far have been with m = 13 or m = 14. A preliminary investigation of resonance with non-zonal harmonics has been made in Ref. 2 and will be summarized here.

When the equations of motion of a satellite in the presence of an arbitrary gravity harmonic are written down and linearized, the result is a set of three equations of the standard form for a forced linear oscillator, one equation for each coordinate. Many frequencies occur in the forcing terms; those which have the form

$$(k\bar{n} - m\omega_e)/2\pi$$

are the ones that may lead to near-resonance. In this, $\bar{n}$ is the satellite mean motion and ω_e is the angular velocity of the earth. (This form is for polar satellites; for non-polar satellites the frequencies also depend slightly upon the precession rates of the mode and of perigee.) k is an integer that depends upon the coordinate in question and upon the parity n of the degree of the harmonic involved.

Near-resonance occurs whenever one of these frequencies is nearly zero. The lowest values of m for which resonance can occur correspond to the case k = 1, and the m values involved in the orbits used to date are 13 and 14. For k = 1, the along-track coordinate of the satellite is the one most sensitive to resonance. Figure IV.2=1 shows the period associated with k = 1 as a function of satellite altitude for five different values of m for circular polar orbits. Note that resonance occurs for smaller values of m as the altitude (and hence the satellite period) increases.

Consider a specific example of a satellite at an altitude of 592 n.m. From Fig. IV.2-1 we see that harmonics with m = 13 contain terms with a period of 2 1/2 days. Any harmonic with odd $n \geq 13$ and m = 13 can produce an along-track oscillation with this period. The amplitude of the oscillation depends both upon the degree n and the size of the harmonic coefficient. For the coefficient with n = 13, the amplitude is 90 meters if the coefficient equals 10^{-6}. Increasing the altitude to 644 n.m. increases the period to 7 days and increases the amplitude to 460 meters for the same coefficient.

This example is for polar orbits. Resonant orbits can occur for any inclination, but the periods for a given altitude change slightly with inclination due to changes in the nodal and perigee precession rates. The amplitudes for given n and m change rapidly with inclination. Therefore if we have several orbits at different inclinations exhibiting resonance with the same value of m we can infer several coefficients with this m and with different n. Unfortunately the amplitudes become very small for inclinations below about 50° unless the resonance is exceedingly close, so that there is a practical limit to the number of coefficients that one can find by the resonance technique. It does appear practical to find the coefficients marked with x in the following scheme from their resonance effects, by using satellites at various altitudes and inclinations:

n \ m	10	11	12	13	14
10	x				
11	x	x			
12		x	x		
13	x	x	x	x	
14				x	x
15			x	x	x
16					
17					x

Resonance with $m > 14$ requires small satellite periods and hence low altitudes. Prospects for tracking such satellites with enough accuracy are discouraging, in view of the large drag that they would encounter. Resonance with $m < 10$ requires high altitude satellites. Such satellites are feasible and can certainly be tracked with accuracy. However the amplitude of the resonance oscillations goes down with increasing altitude and extremely close resonance is needed in order to give a measurable perturbation. Except by accident, or in exceptional circumstances, it is unlikely that high altitude satellites can be used to study resonance. One exceptional circumstance occurs with the synchronous satellites, whose periods are purposely and carefully adjusted to agree closely with the length of the day and which are therefore in extremely close resonance.

Resonance can also occur with the value of $k = 2$ in the expression for the resonance period. This type of resonance occurs with m in the range

26 - 28 for the satellites used so far, and has not been studied in detail.

At the beginning of this section, we pointed out that three parameters, the semi-major axis, the eccentricity, and the inclination have a significant effect upon the influence coefficients of a gravity harmonic. So far we have discussed almost entirely the effect of inclination. Now we turn to the eccentricity.

The effect of eccentricity is to destroy the symmetry that exists in the effect of a harmonic upon a circular orbit. The force produced by a harmonic has three components which are proportional to the three components of the vector gradient of the harmonic in the potential function. Let us resolve this force into its vertical component and into two horizontal components, one in the plane of the orbit and one normal to it. Use the term "forward component" to denote the horizontal component in the plane of the orbit and the term "cross component" to denote the other horizontal one.

Let us now inspect these three components for two points at the same radius from the center of the earth but diametrically opposed; that is, let us inspect the inversion symmetry of the forces. We find that each component is either unchanged or simply changes sign. It never happens that all three components arising from a given harmonic have the same behavior. Typically, if the forward component changes sign, the other two do not, and vice versa.

Now we must consider what happens to a satellite in a circular orbit at two points separated by half a revolution. If the earth did not rotate, this would be determined only by the inversion symmetry; since the earth does rotate we must consider how the force field of a given harmonic has rotated during half a satellite revolution. The rotation of the earth is not important for small m, but becomes crucial at values of m large enough for resonance to occur.

We still find that the three components of the force tend to have different symmetry behavior. For small m, for which the rotation does not matter, the symmetry behavior depends only upon n; for large m, the symmetry depends upon both m and n.

If a force component changes sign on opposite sides of the orbit, it tends to cancel for an entire revolution and the corresponding perturbation is small. If the component does not change sign, the perturbation is large. Thus, on the average, only half of the coordinates of a satellite respond in a significant manner to a given harmonic, for a circular orbit.

If the orbit has sensible eccentricity, the satellite is at a different radius on opposite sides of the orbit (except for the points close to the ends of the semi-minor axis), hence the sizes of the components are necessarily different and cancellation cannot occur. Those components which do not cancel for circular orbits are hardly affected by a moderate eccentricity. Those which do cancel for circular orbits give perturbations proportional to the eccentricity for moderate eccentricity.

On the average, then, twice as many satellite coordinates show perturbations for eccentric orbits as for circular orbits. Thus we would expect twice the independent information and the ability to infer twice as many harmonic coefficients from eccentric orbits as from circular orbits.

This statement would probably be true if we could vary the eccentricity of an orbit independently of the other parameters. Unfortunately we cannot. The minimum perigee altitude of a satellite must be set so that it does not encounter the earth's crust nor indeed too much of its atmosphere in order to have a useful lifetime. We can increase the eccentricity only by increasing the apogee altitude and the average altitude at the same time. As we increase the average altitude we decrease the force due to a harmonic at a high rate. The perturbation due to a harmonic varies inversely as the semi-major axis raised to the power $n + (3/2)$; for large n this cuts the perturbation drastically for higher altitudes.

Overall, we favor having some satellites with a moderate eccentricity, say in the range 0.02 to 0.04, primarily to give a better definition to the position of perigee and hence to improve the analysis of the zonal harmonics, and secondarily to give some more information. Large eccentricities are not favored unless they are a part of the practical problem of obtaining a large semi-major axis.

There are two advantages in having a large semi-major axis for some of the satellites used in studying the gravity field. First, as long as we rely only upon variation in inclination to separate harmonics, we begin to run into trouble when the number of inclinations becomes large and their separation necessarily becomes small. Changing the semi-major axis also changes the relative effects of harmonics. To be more specific, the effect of a harmonic falls off rapidly with increasing degree n, as we mentioned above. This strongly separates harmonics of low and high degree. In effect, the high altitude satellites give us the low degree harmonics with little contamination from those of high degree; the low altitudes then give us more information about harmonics of high degree.

Second, high altitude satellites enable us to find the harmonic J_0, more commonly denoted by K, which is the product of the total mass of the earth and the gravitational constant. In all determinations of the geopotential from satellite data that have yet been made, it has been found that K and the mean equatorial radius R_e cannot be well determined. The reason for this is that the satellite equations of motion can be written to high accuracy in an invariant form in terms of dimensionless variables. In these dimensionless variables, the unit of length is R_e and the unit of time is $\sqrt{R_e^3/K}$. From the data, time relations and hence K/R_e^3 are well determined, but there is little sensitivity to distance relations and hence R_e and K individually are poorly determined. The reason that there is little sensitivity to distance is that we can change the average radius of a satellite orbit and compensate this change almost exactly by changing the radii of the stations.

This near redundancy in the two constants K and R_e has been studied in Ref. 3 and it was found that they could be determined independently by using at least one orbit with a markedly different altitude. The study was based upon a simulation to determine quantitatively the sensitivity to errors in K and R_e (using R_e as the scaling constant for station radii). In the simulation the following conditions were assumed:

(1) Sufficient satellites at differing inclinations are available to determine accurately all geodetic constants except K and R_e.

(2) Station radii are initially adjusted to least-squares fit the doppler tracking data over all passes of satellites at 1000 km altitude in the presence of an error in K.

The ability to distinguish between a genuine error in station radius and an apparent error due to an error in K was measured by determining the amount of movement in station radius (relative to the least squares value described above) that was required to minimize doppler data residuals for individual passes as a function of satellite altitude during the pass. The results are summarized in Fig. IV.2-2.

As an example in using Fig. IV.2-2, suppose that we try to infer K and R_e using a satellite at 1000 km. altitude. The curves show position errors in the stations when the error in K is 1 part in 10^5. For a satellite at 1000 km. altitude, the curve marked "error in least square value of station altitude" shows that the error in K is best compensated by an RMS error of about 12 meters in the station radii. The other curve shows that the residuals (the noise level) are increased by about 1.5 meters. At present, the noise level is about 75 meters, and this change in noise level produced by an erroneous K is not detectable, and will hardly be detectable even when the noise level is decreased to 5 meters by having enough satellites of differing inclinations at low altitude available. When we achieve a sensitivity of about 5 meters, the curve shows us that we can detect an error of 1 in 10^5 in K; in fact, we can probably detect about half this amount. If we can find K to 5 in 10^6, we can find R_e to 1/3 the relative error or to about 10 meters.

Tests with actual data have shown that we can do better than Fig. IV.2-2 indicates and that even in the present situation we can detect about 5 parts in 10^5 in K or about 100 meters in R_e. With the use of a satellite having a maximum altitude of 2500 km., we can thus expect to find R_e within 1 or 2 meters. It is not necessary that the average altitude be 2500 km., and it is perfectly acceptable to use an eccentric orbit to achieve this maximum altitude if this eases vehicle or payload problems.

In summary, extending the gravity analysis through harmonics of degree 16, and finding K and R_e, requires about four orbits substantially different from those now available. One of the satellites should differ by having a maximum altitude of about 2500 km.; for the others, varying the inclination is probably sufficient. A small eccentricity, in the range 0.02 - 0.04 is recommended for at least some of these.

References - Section IV.2

1. Newton, R. R., Geodesy by Satellite, Science 144, no. 1620, pp. 803-808, May 15, 1964.

2. Yionoulis, S. M., Resonant Geodesy, Applied Physics Lab. Report TG-633, December 1964.

3. Guier, W. H., Note on Determining the Overall Distance Scale of the Earth Using High Altitude Satellites, Applied Physics Lab. Report CF-3098, July 1964.

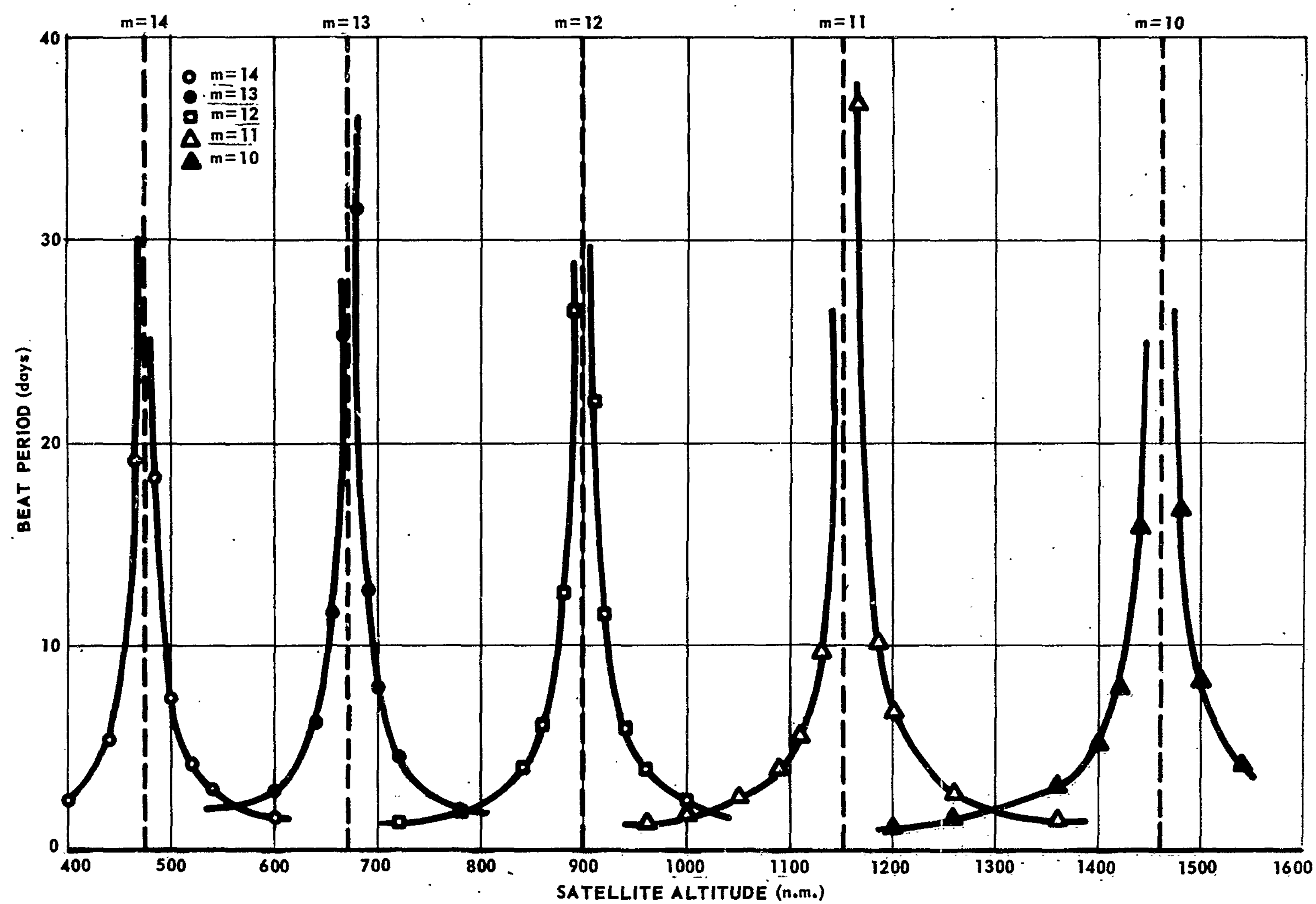

Fig. IV.2–1. BEAT-PERIOD ARISING FROM NEAR RESONANCE, AS A FUNCTION OF SATELLITE ALTITUDE, FOR VARIOUS ORDERS m OF HARMONIC.

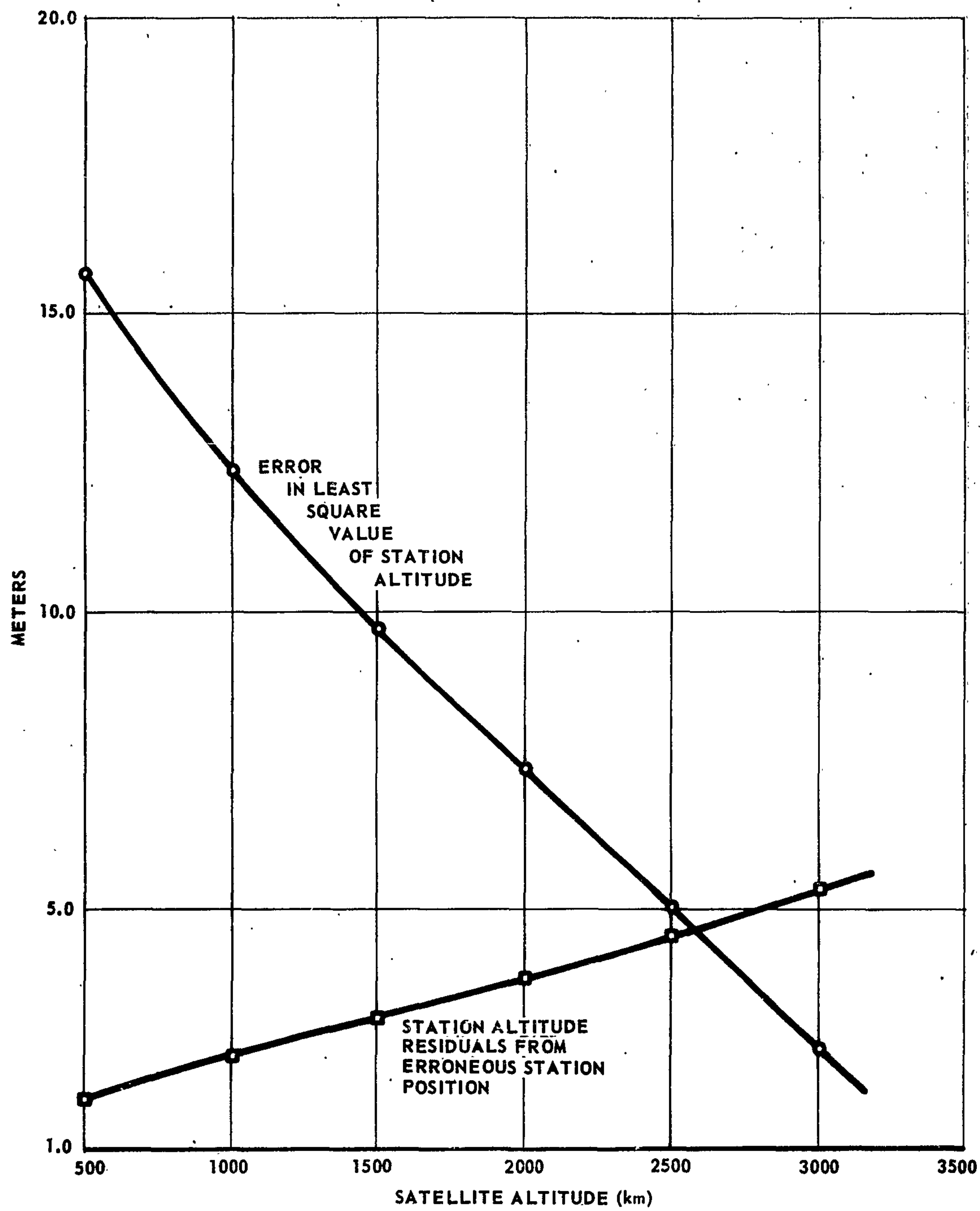

Fig. IV.2-2. SENSITIVITY OF DETERMINING K AND R_e AS A FUNCTION OF SATELLITE ALTITUDE, DRAWN FOR AN ERROR OF $1/10^5$ IN K.

The Johns Hopkins University
APPLIED PHYSICS LABORATORY
Silver Spring, Maryland

INITIAL DISTRIBUTION - APPLIED PHYSICS LABORATORY

Newton, R. R. (10)
Kershner, R. B.
Riblet, H.
Ellis, R.
Worsley, G.
Wyatt, T.
Guier, W. H. (2)
Weiffenbach, G. C. (2)
Distribution Project (50)

INITIAL DISTRIBUTION EXTERNAL TO APL

Chief, Bureau of Naval Weapons
Department of the Navy
Washington 25, D. C.
Attn: Lt. Cdr. Limerick/RTMS (50)
DLI-31 (2)

Bureau of Naval Weapons Representative
APL/JHU
8621 Georgia Avenue
Silver Spring, Maryland

Defense Documentation Center for
Scientific and Technical Information
Building 5, Cameron Station
Alexandria, Virginia
Via: Bu-Weps DLI-302 (20)